100

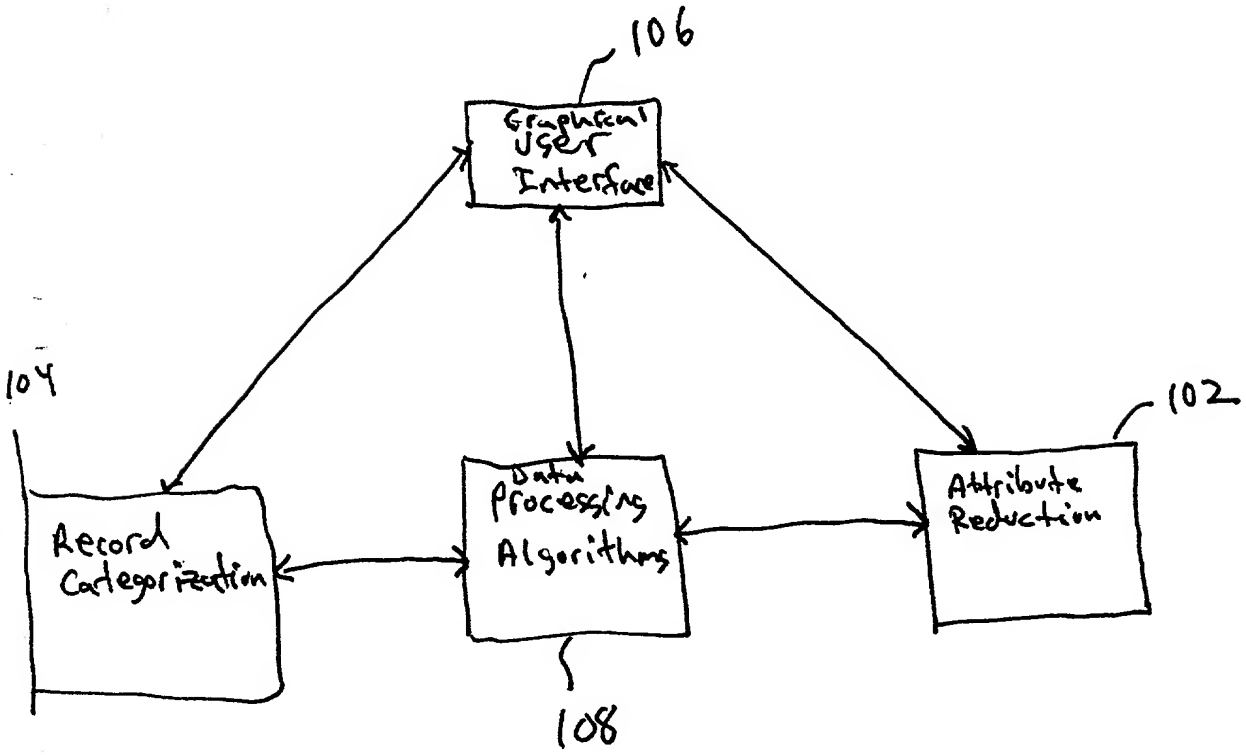


Figure 1

EV012824282US

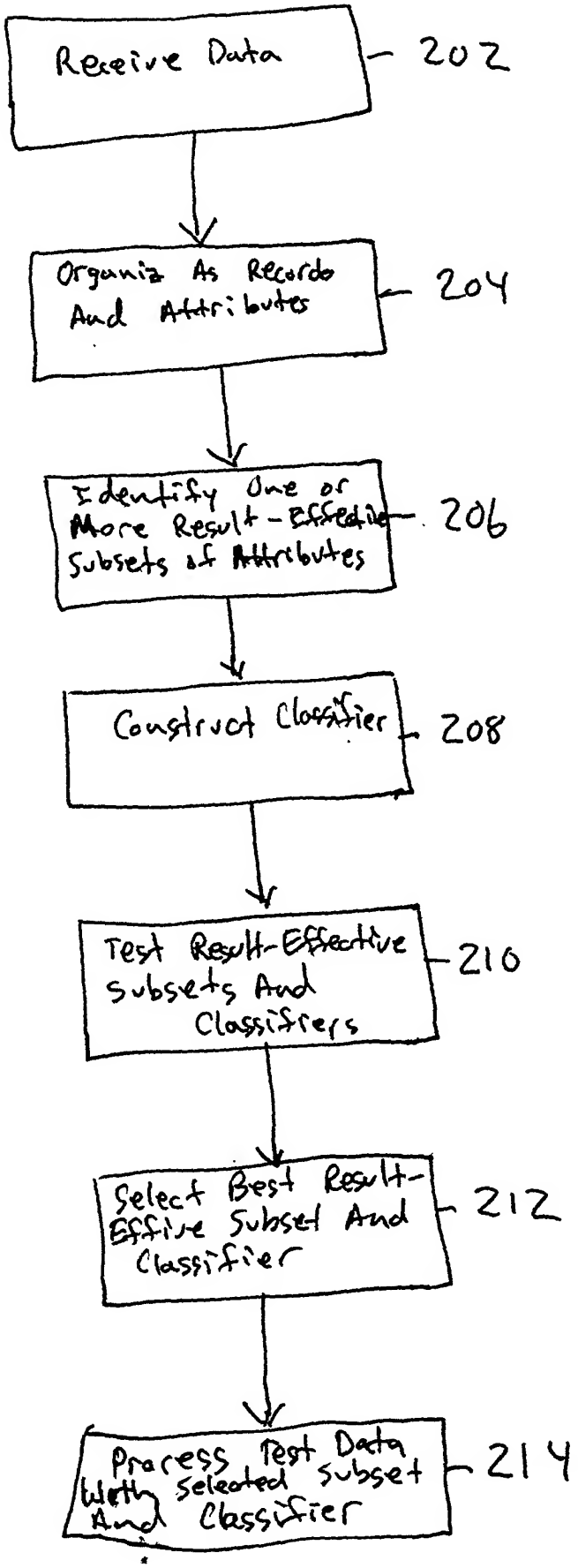

Receive Data
202
200
Organiz As Records And Attributes
204
Identify One or More Result-Effective Subsets of Attributes
206
Construct Classifier
208
Test Result-Effective Subsets And Classifiers
210
Select Best Result-Effive Subset And Classifier
212
Process Test Data With Selected Subset And Classifier
214

Figure 2

221897

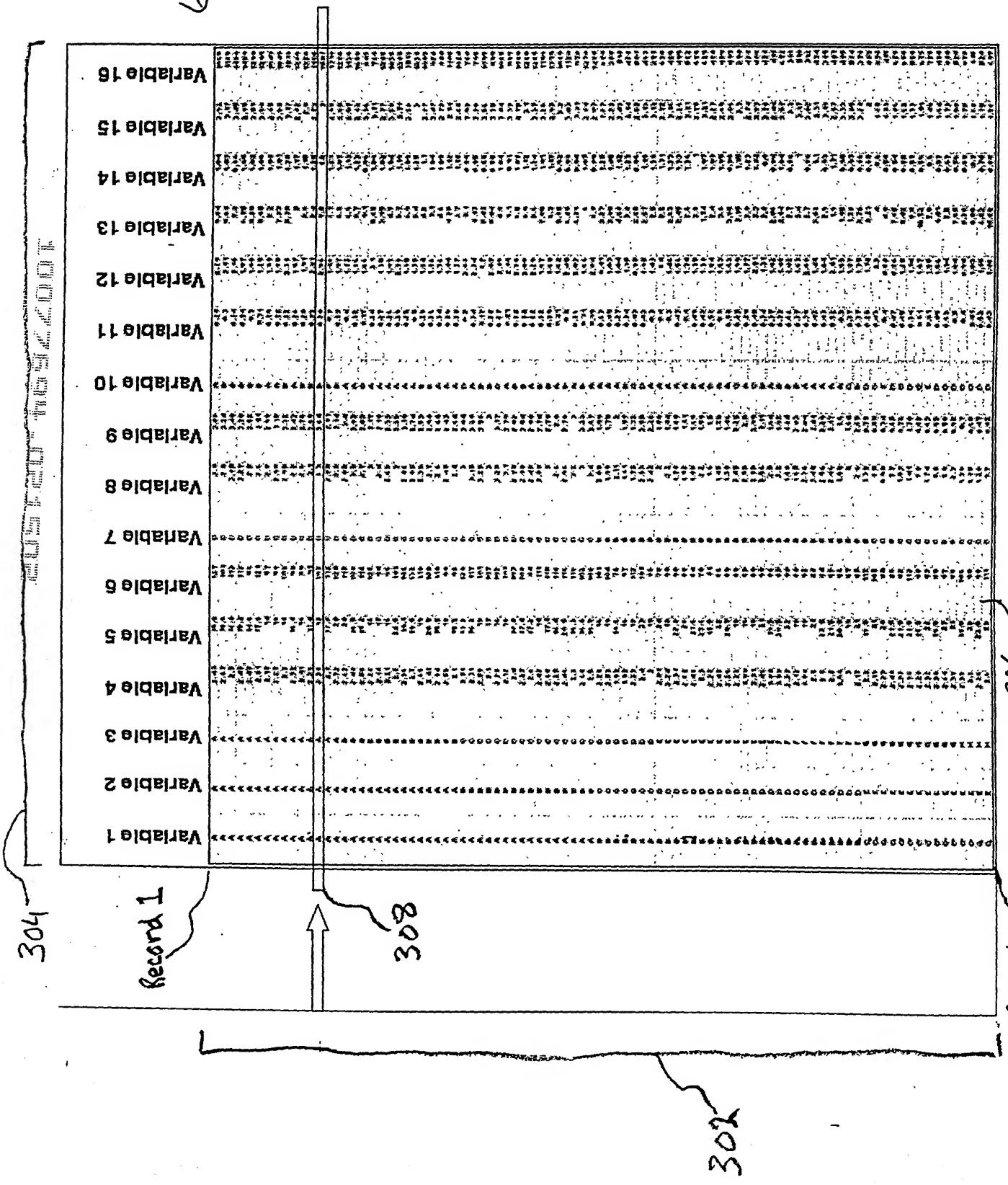


Figure 3

20STEP "1692200F

304

Variable 4

400

404

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 9	Variable 10	Variable 11	Variable 12	Variable 13	Variable 14	Variable 15	Variable 16
B	D	F	[illegible]	20.80	86	B	1.98	1.60	A	0.30	1.53	1.95	0.95	3.33	495
A	A	B	2.21	16.00	96	D	2.65	2.33	A	0.26	1.98	4.70	1.04	3.59	1035
B	D	F	[illegible]	19.00	92	B	2.36	2.04	A	0.39	2.08	2.70	0.86	3.02	312
A	A	A	2.67	18.60	101	C	2.80	3.24	B	0.30	2.81	5.68	1.03	3.17	1185
A	A	B	2.70	22.50	101	C	3.00	3.25	A	0.29	2.38	5.70	1.19	2.71	1285
B	D	E	2.42	22.00	86	B	1.45	1.25	A	0.50	1.63	3.60	1.05	2.65	450
B	D	F	2.02	19.00	80	B	1.65	2.03	B	0.37	1.63	3.40	1.00	3.17	510
A	A	B	2.55	18.00	98	C	2.45	2.43	A	0.29	1.44	4.25	1.12	2.51	1105
A	B	C	2.60	17.20	94	C	2.45	2.99	A	0.22	2.29	5.60	1.24	3.37	1265
B	D	F	2.17	21.00	85	B	2.60	2.65	B	0.37	1.35	2.76	0.86	3.28	378

308

306

402

LOW HIGH

402a 402b 402c

Figure 4A

2218917

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 4 of 79

20ST20 th9z20ot

304

Variable 2

400

406

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 9	Variable 10	Variable 11	Variable 12	Variable 13	Variable 14	Variable 15	Variable 16
B	D	F	1.99	20.80	86	B	1.98	1.60	A	0.30	1.53	1.95	0.95	3.33	495
A		B	2.21	16.00	96	D	2.65	2.33	A	0.26	1.98	4.70	1.04	3.59	1035
B	D	F	1.94	19.00	92	B	2.36	2.04	A	0.39	2.08	2.70	0.86	3.02	312
A		A	2.67	18.60	101	C	2.80	3.24	B	0.30	2.81	5.68	1.03	3.17	1185
A		B	2.70	22.50	101	C	3.00	3.25	A	0.29	2.38	5.70	1.19	2.71	1285
B	D	E	2.42	22.00	86	B	1.45	1.25	A	0.50	1.63	3.60	1.05	2.65	450
B	D	F	2.12	19.00	80	B	1.65	2.03	B	0.37	1.63	3.40	1.00	3.17	510
A		B	2.55	18.00	98	C	2.45	2.43	A	0.29	1.44	4.25	1.12	2.51	1105
A	E	C	2.60	17.20	94	C	2.45	2.99	A	0.22	2.29	5.60	1.24	3.37	1265
B	D	F	2.17	21.00	85	B	2.60	2.65	B	0.37	1.35	2.76	0.86	3.28	378

308

306

402

LOW HIGH

402a 402b 402c

Figure 4B

2218917

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 5 of 79

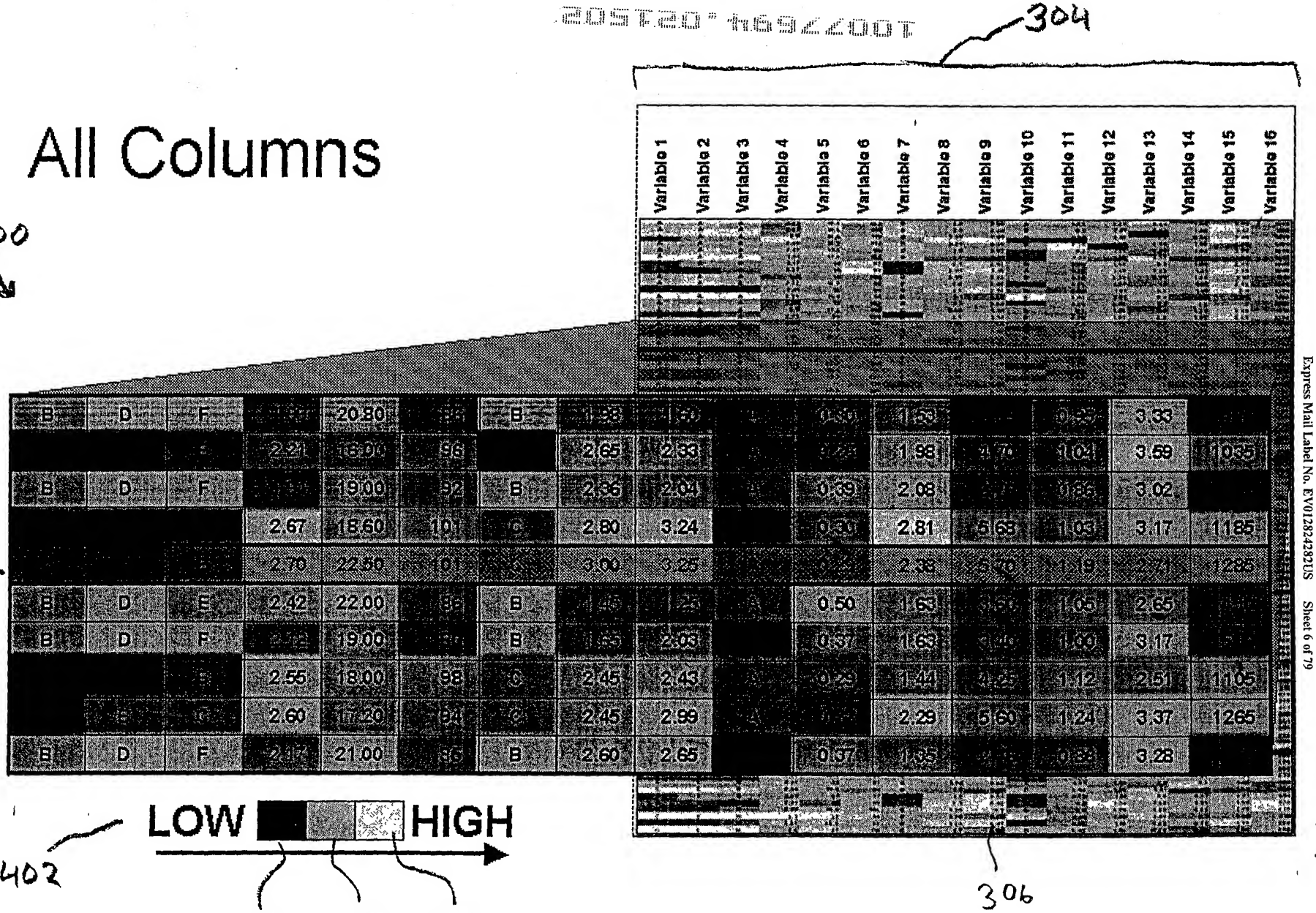

All Columns
400
304
Variable 1
Variable 2
Variable 3
Variable 4
Variable 5
Variable 6
Variable 7
Variable 8
Variable 9
Variable 10
Variable 11
Variable 12
Variable 13
Variable 14
Variable 15
Variable 16
308
306
LOW
HIGH
402
402a
402b
402c

Figure 4C

20ST20 "A692200T

Completed view with numbers removed

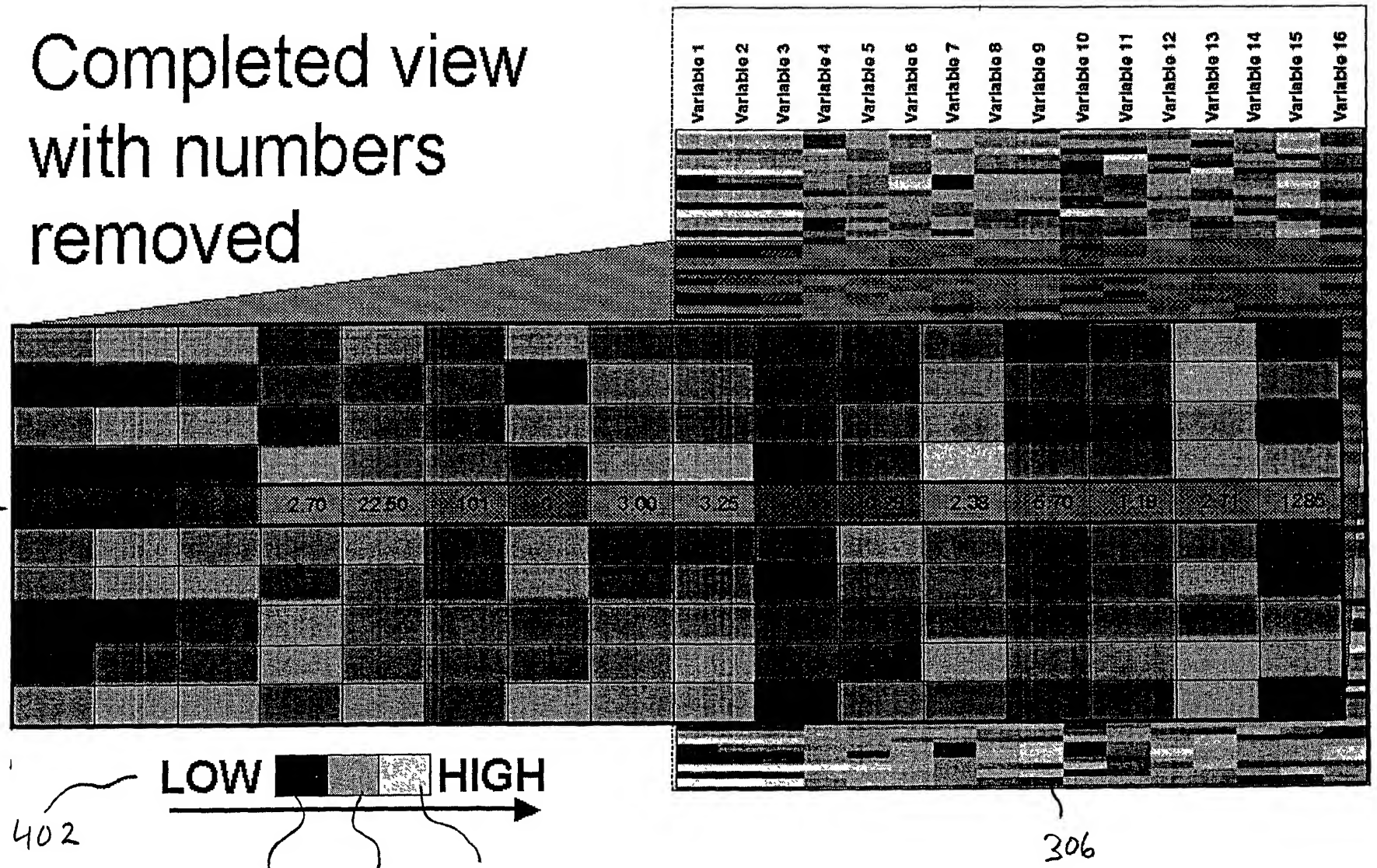


Figure 4D

Ideal Case

500

Random

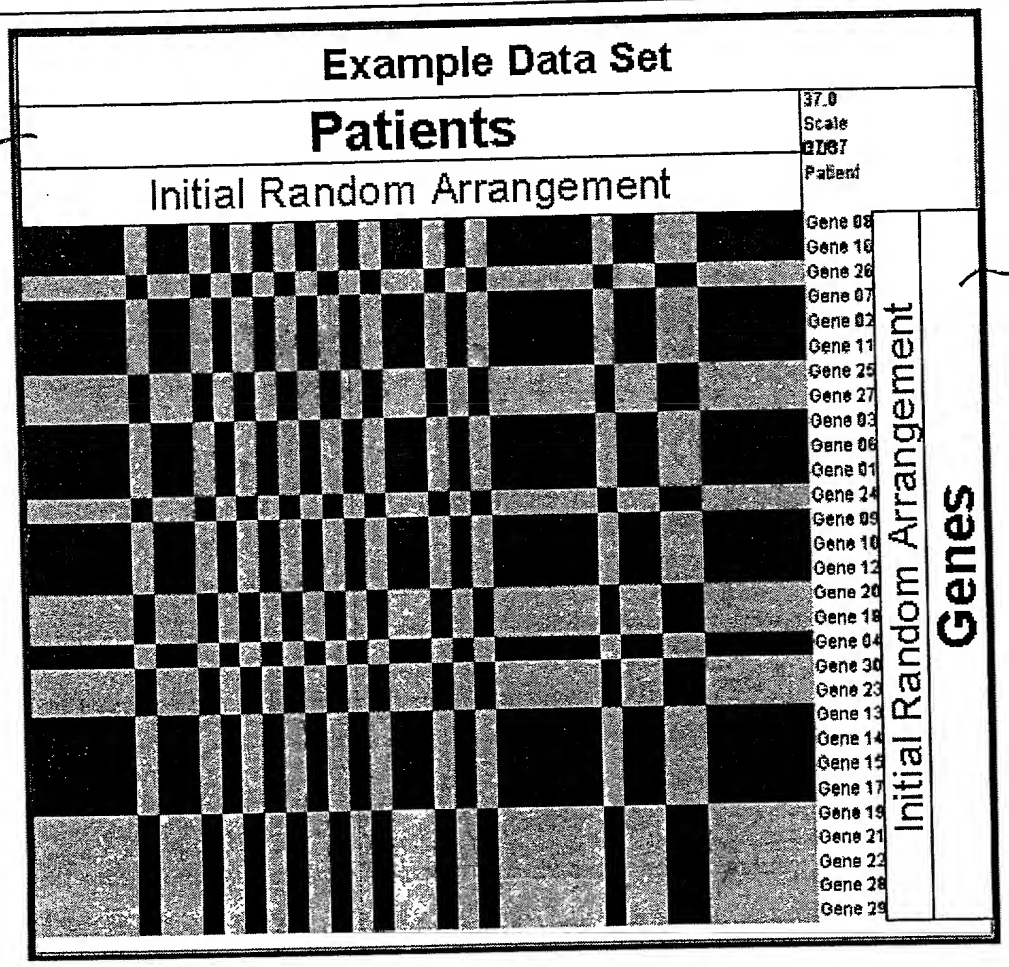


Figure 5A

500

Ideal Case

Sorted by Patients

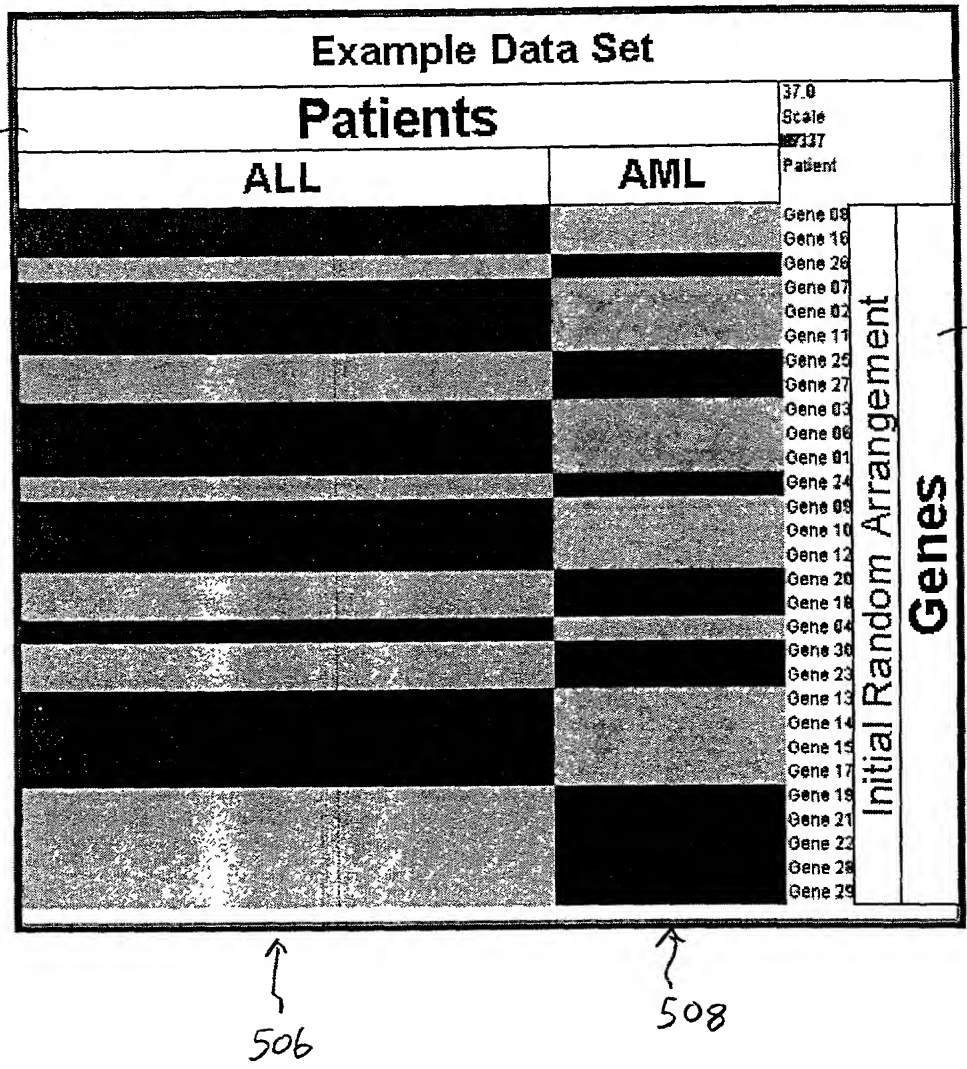


Figure 5B

2218917

Ideal Case

500

Grouped on Genes

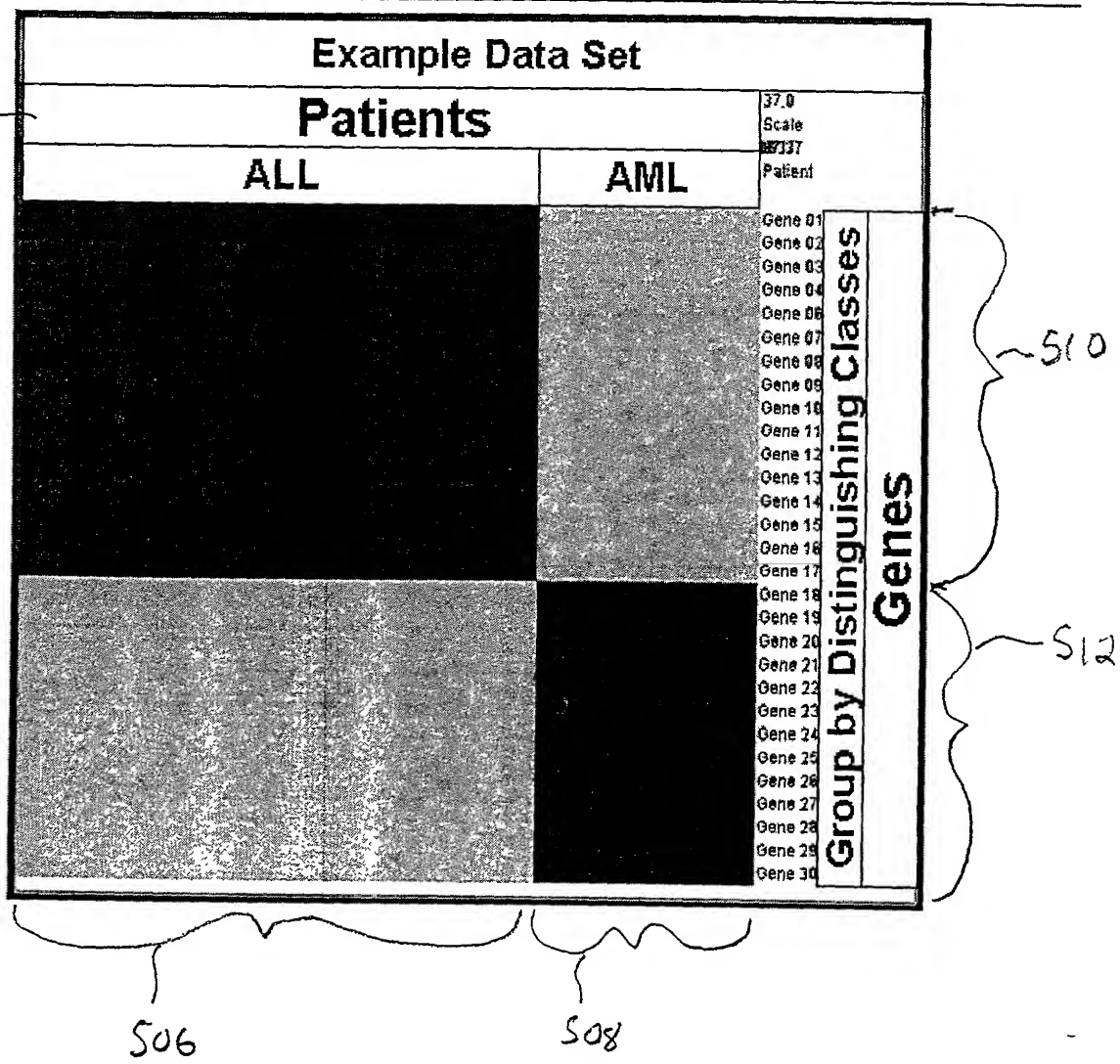


Figure 5C

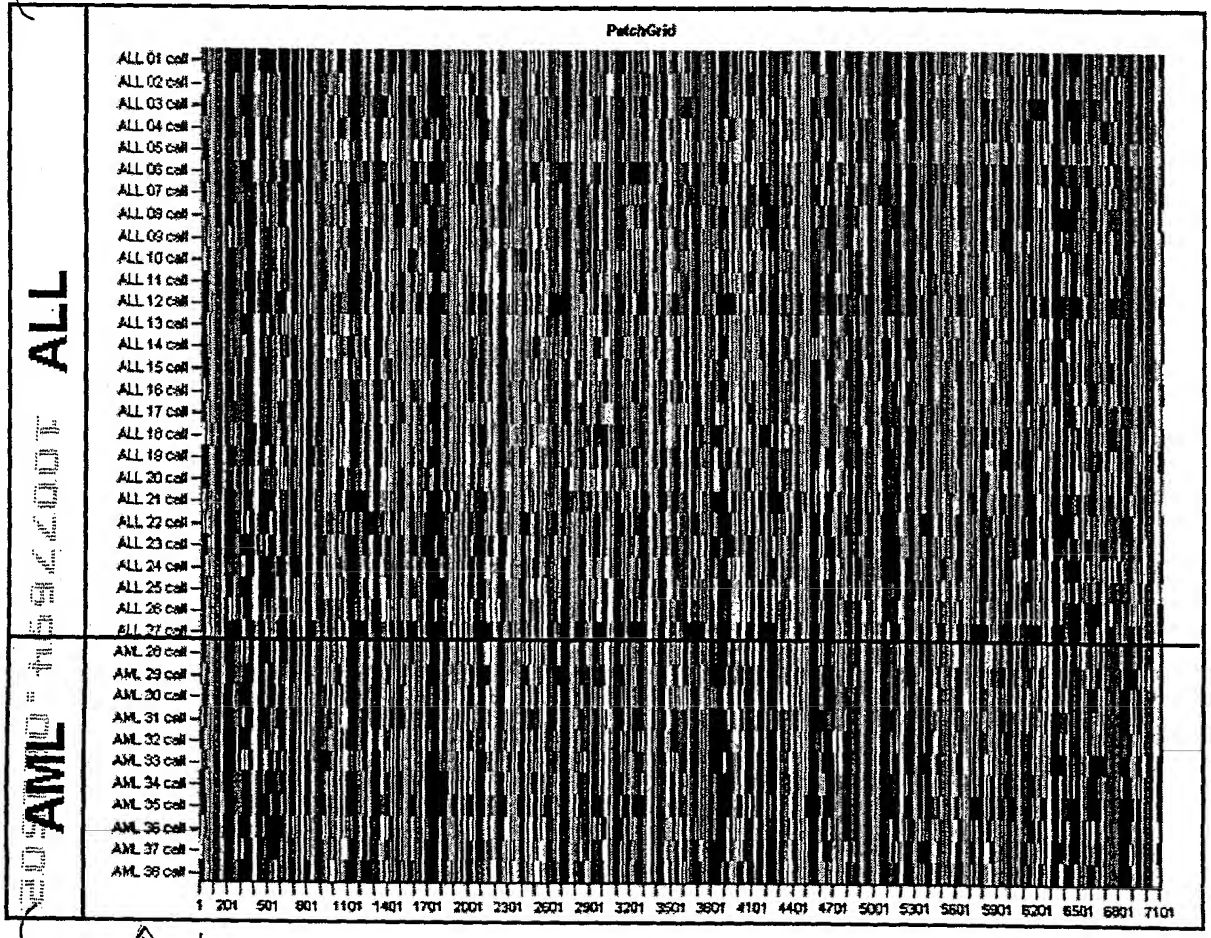
221897

5C

Random

600

606



608

602

604

Figure 6A

2218917

600

Sort 1st Dimension

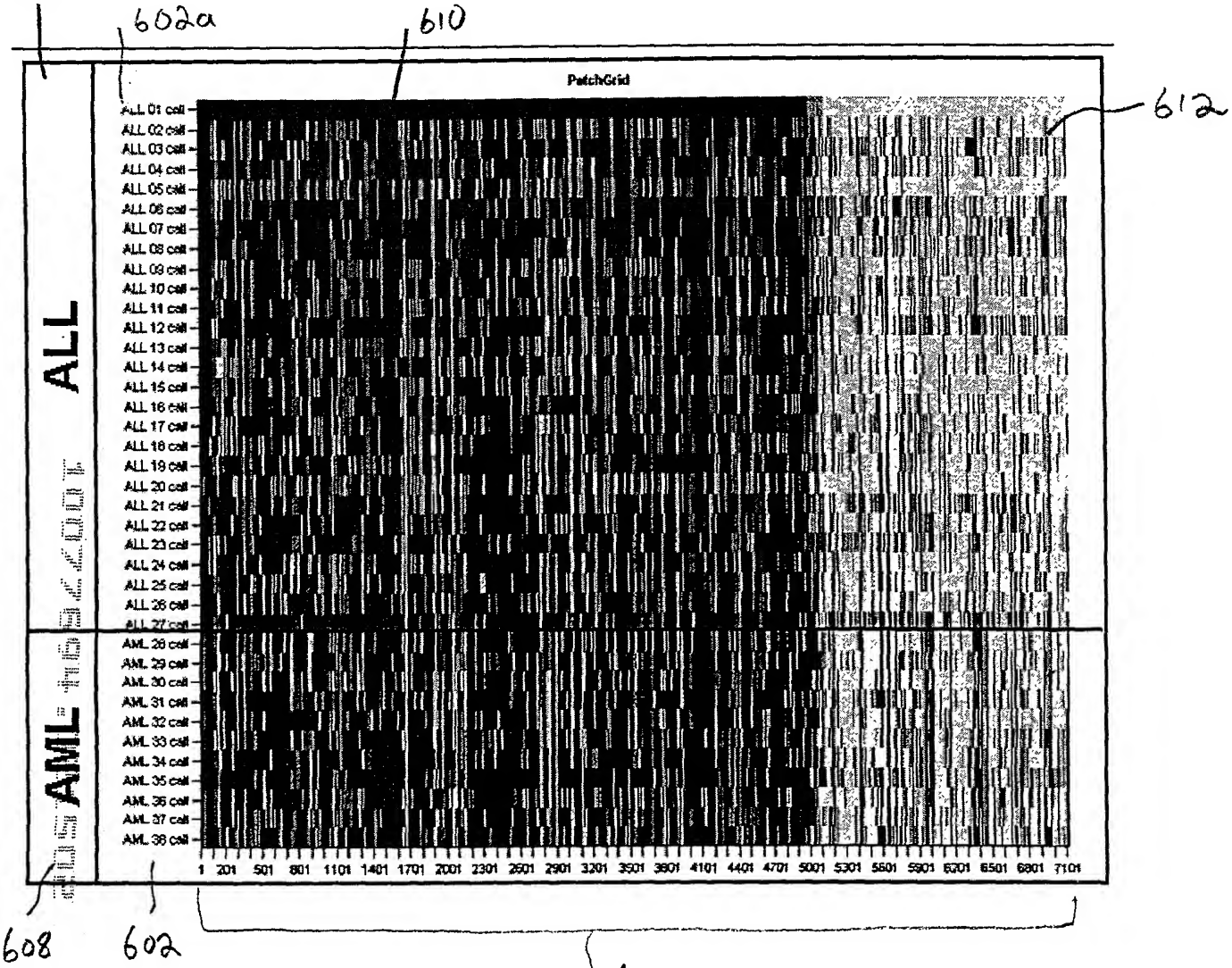


Figure 6B

2218917

600

Sort 2nd Dimension

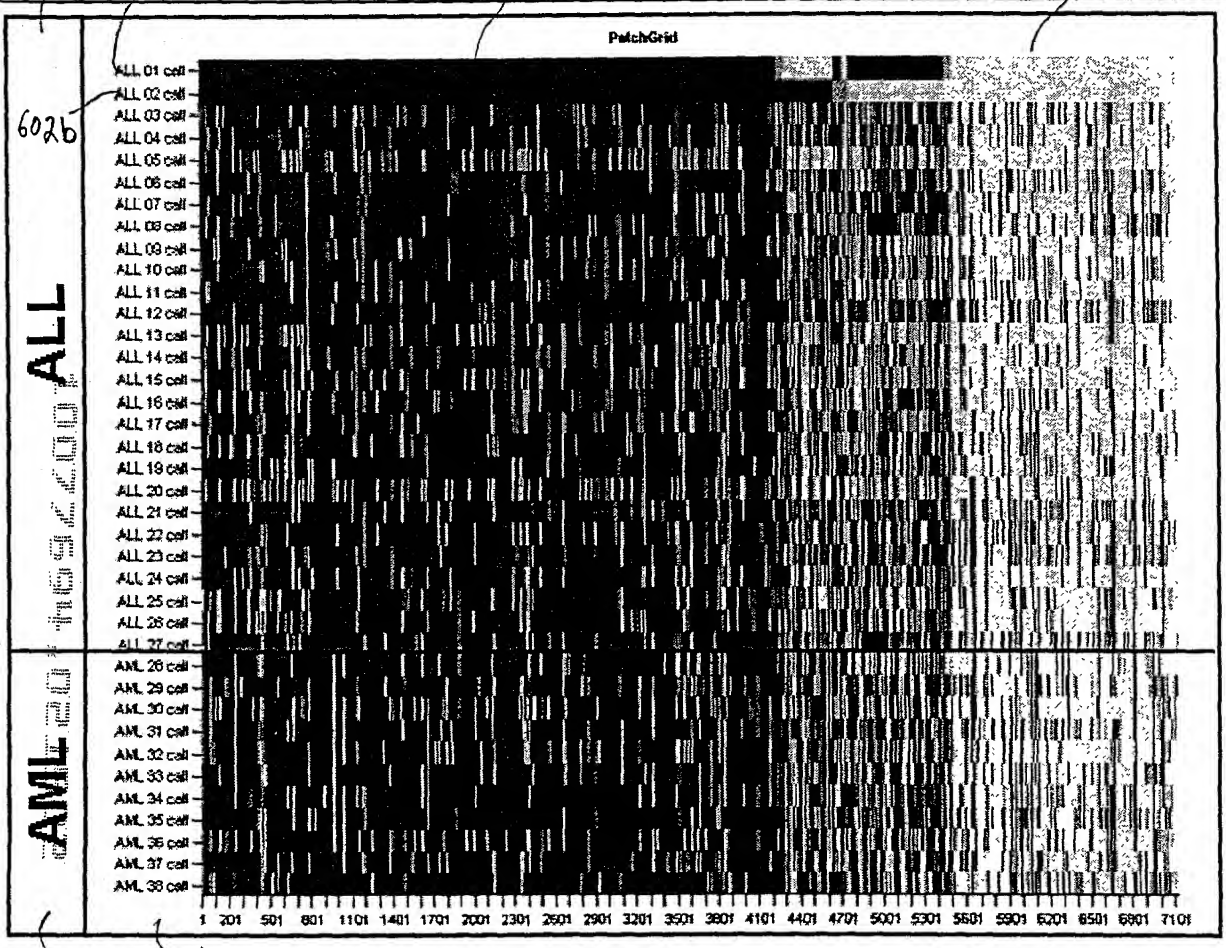


Figure 6C

2218917

6c

Sorted 5 Dimensions

600

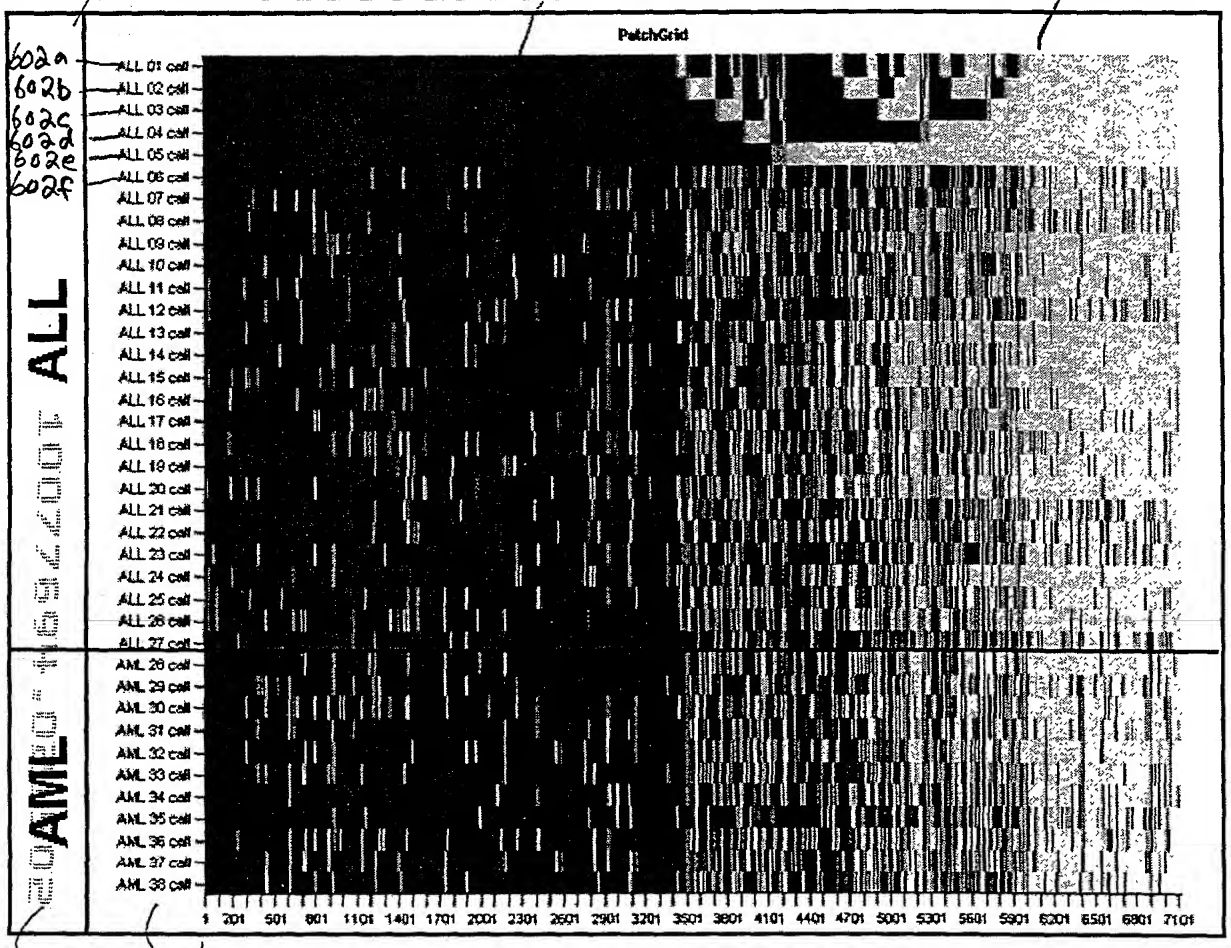


Figure 6D

2218917

6D

Sorted 33 Dimensions

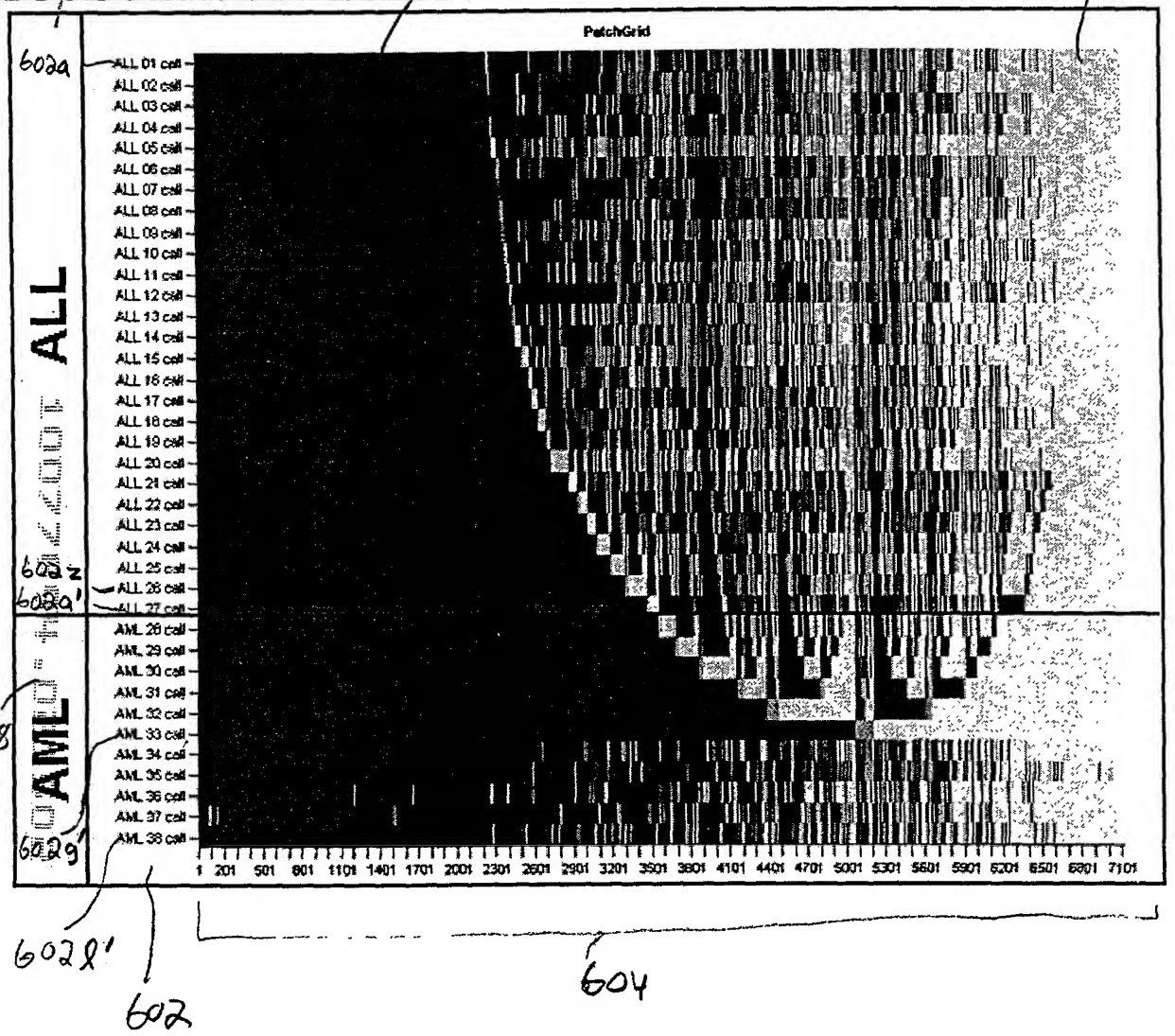


Figure 6E

2218917

All Dimensions Sorted

600

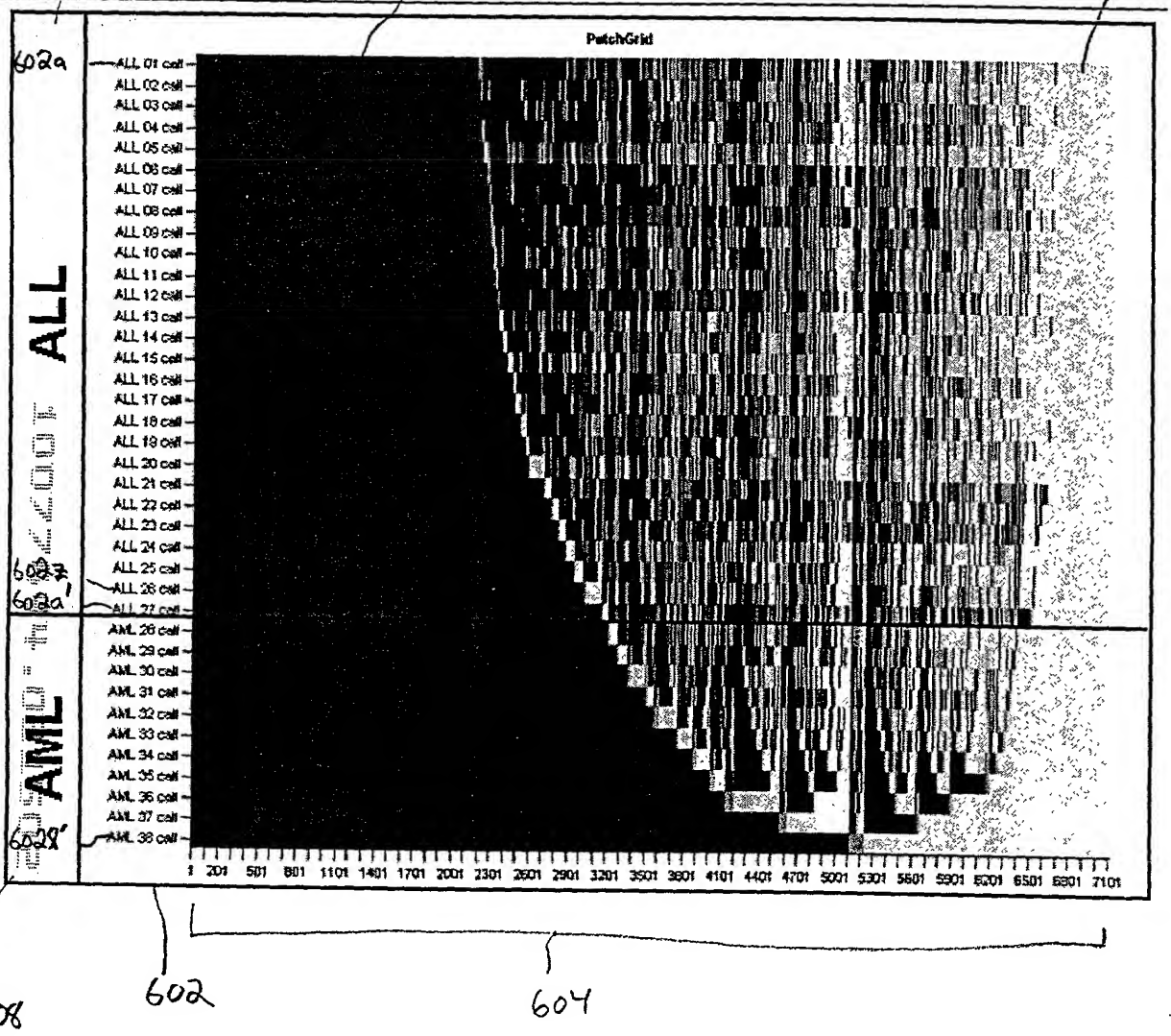


Figure 6F

22189 (7

6F

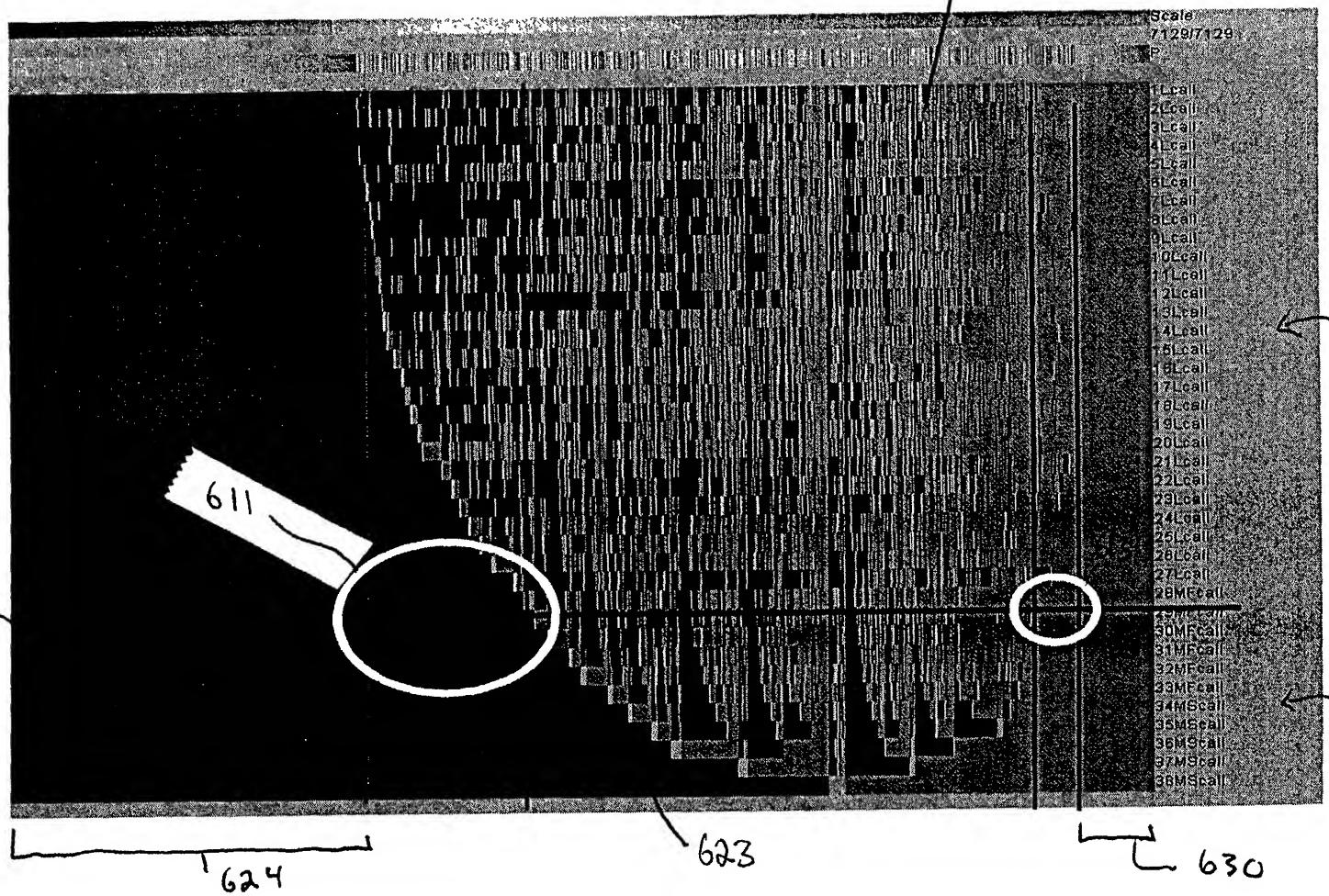


Figure 6G

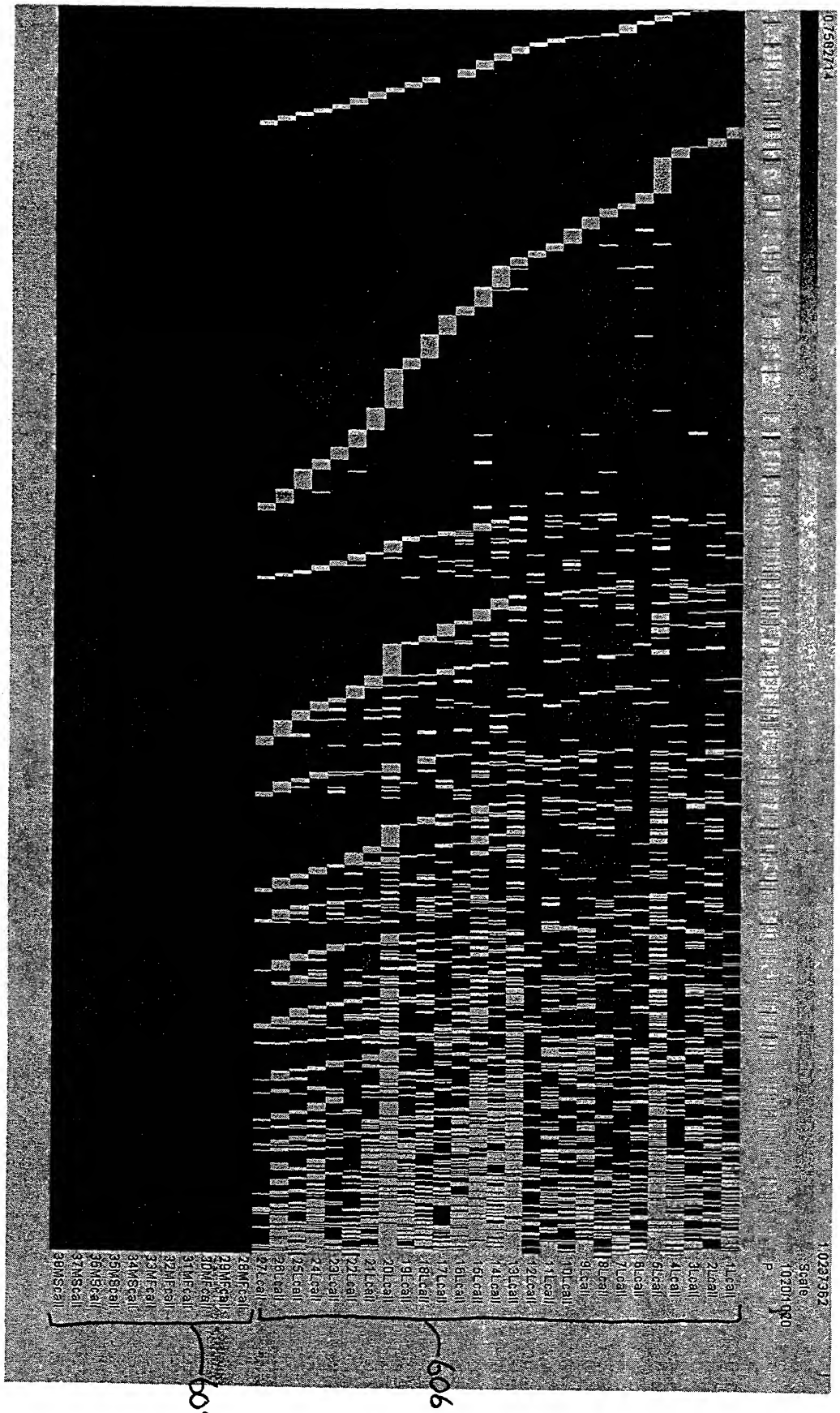


Figure 6B

637

20520 h6926007

Principal Uncorrelated Record Selection

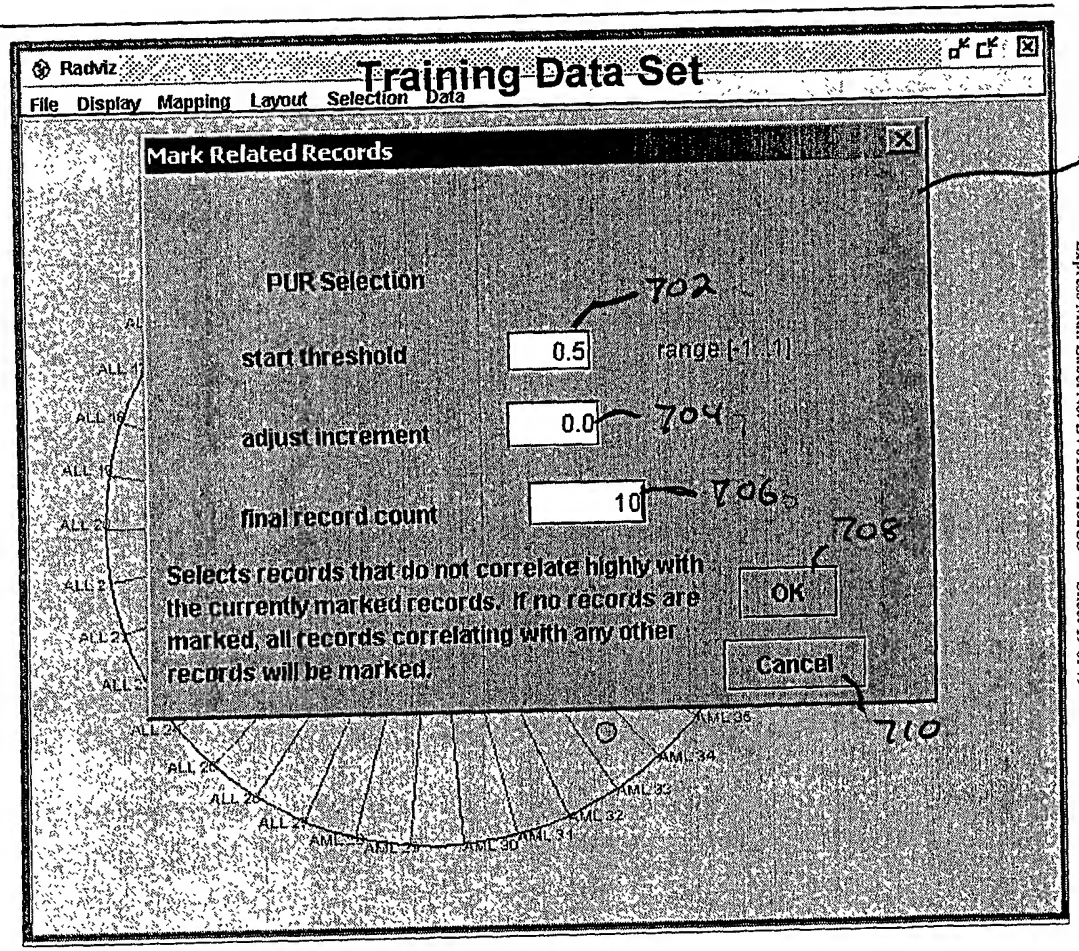


Figure 7

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 19 of 79

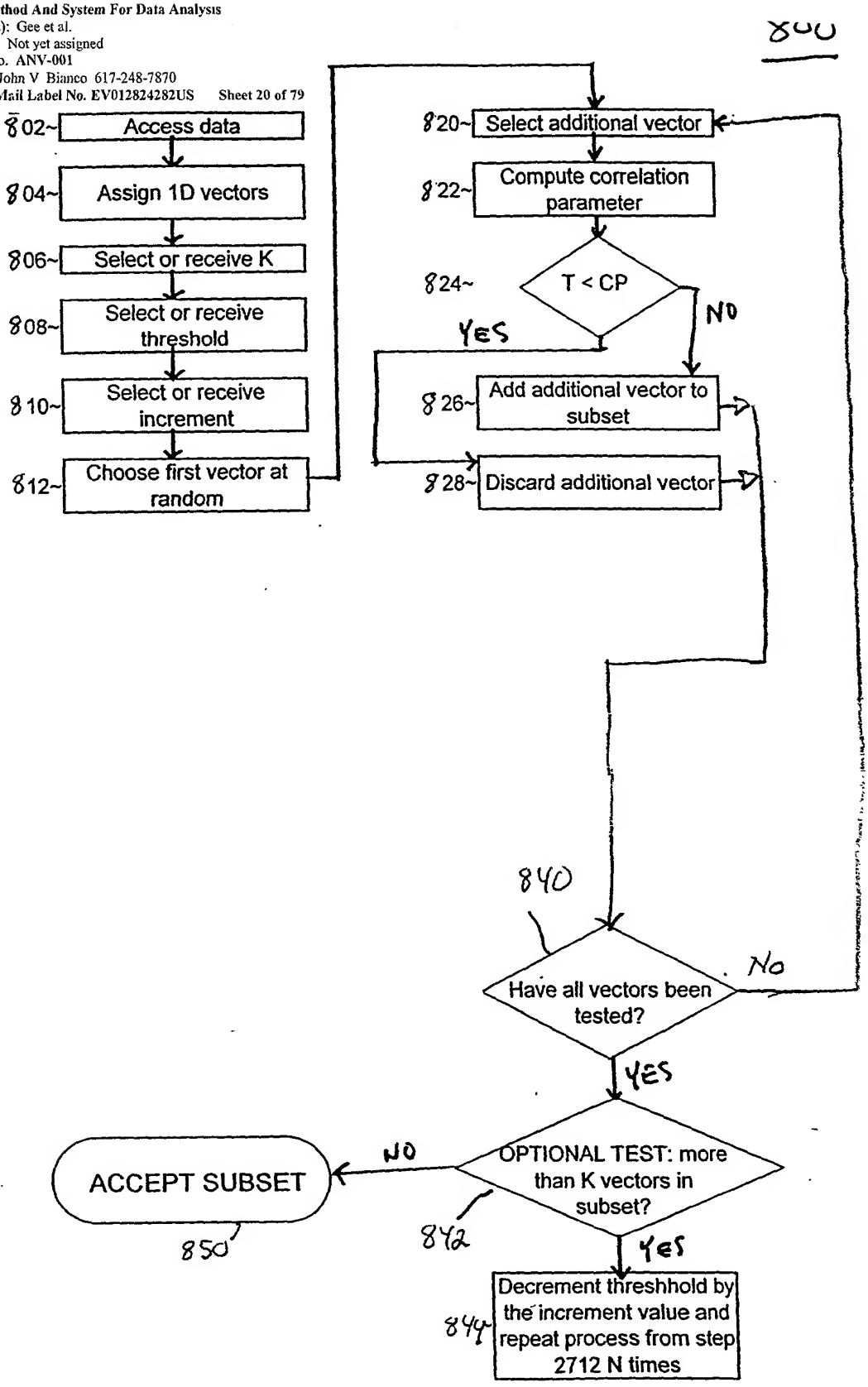


Figure 8.

2268791

76 Principal Uncorrelated Genes

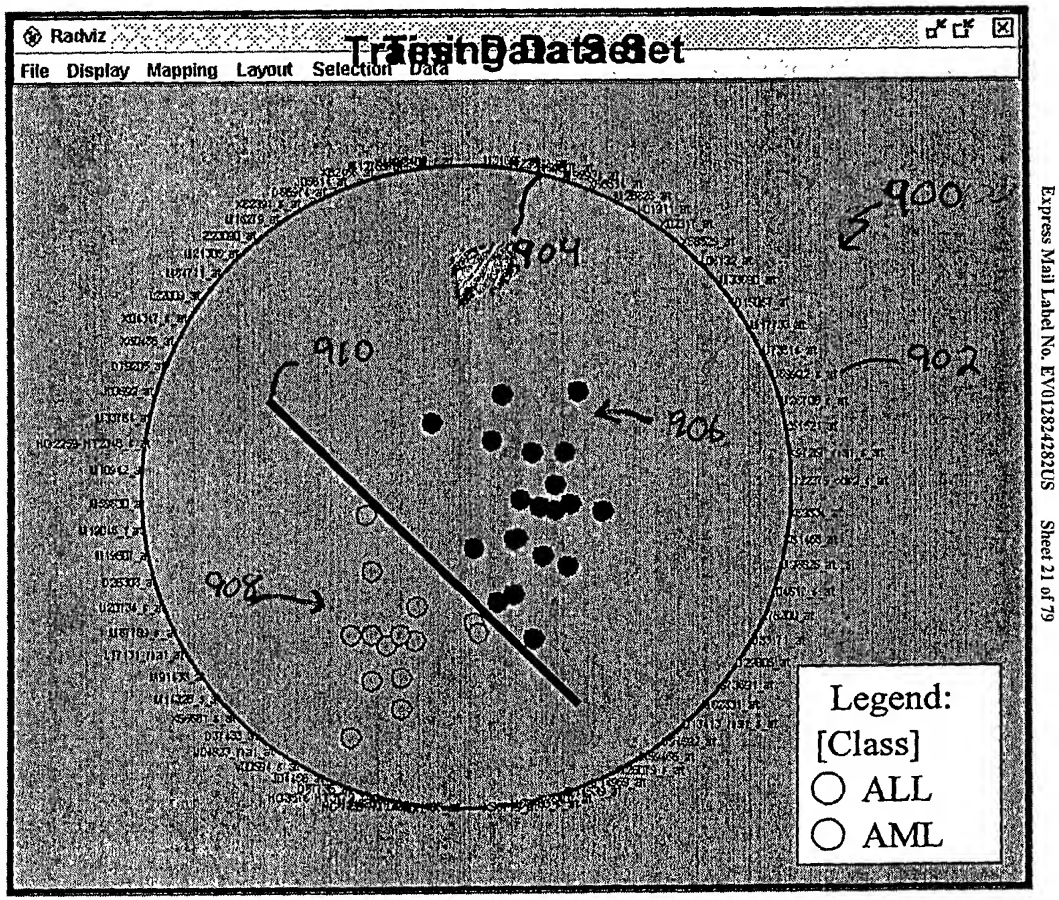


Figure 9

20512018692007

Statistical Metadata 1000

(Affymetrix Absolute Data)

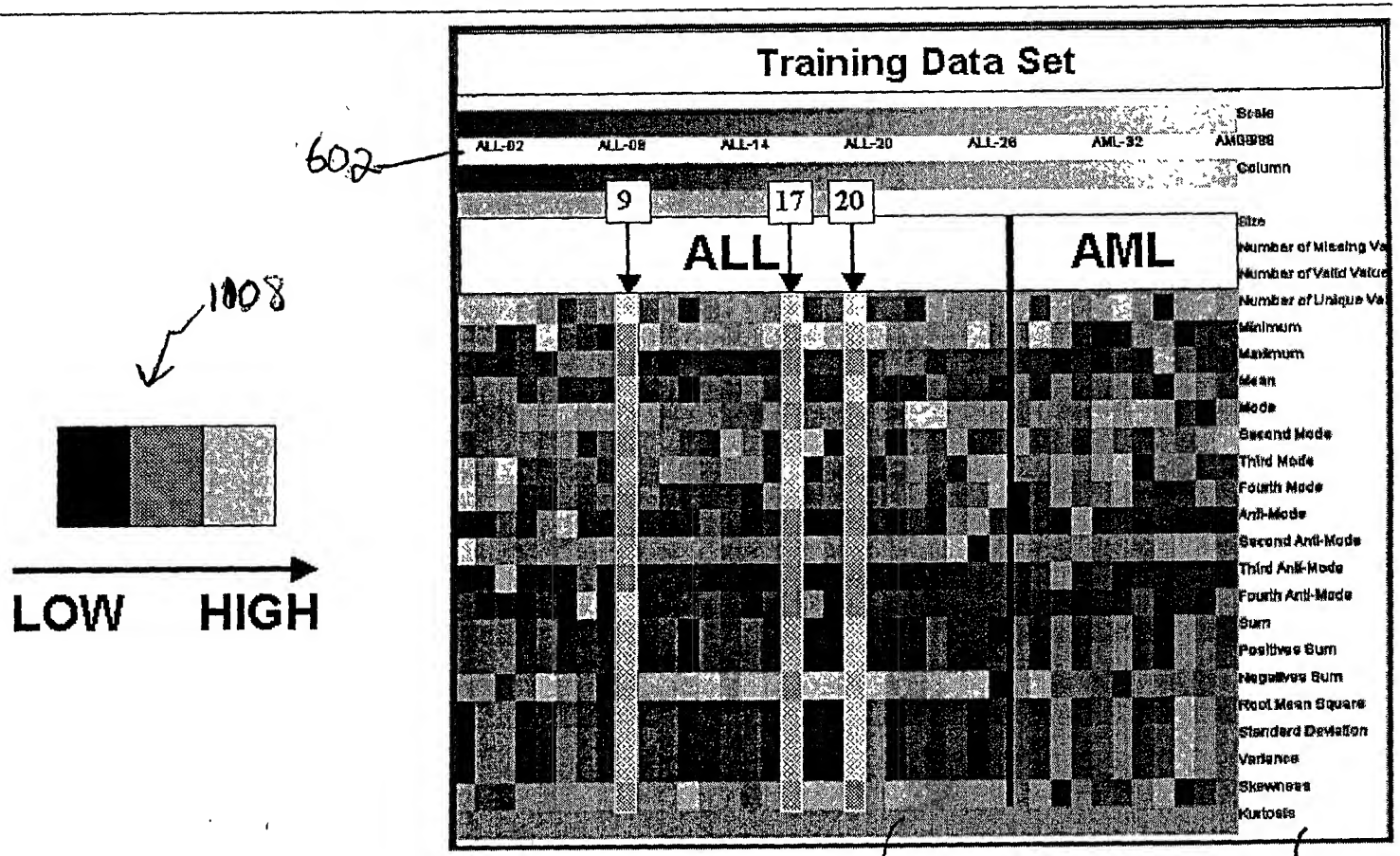


Figure 10A

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012842820US Sheet 22 of 79

20STP20 "h697C00T

Pearson Correlation Coefficients

(Affymetrix Absolute Data)

1001

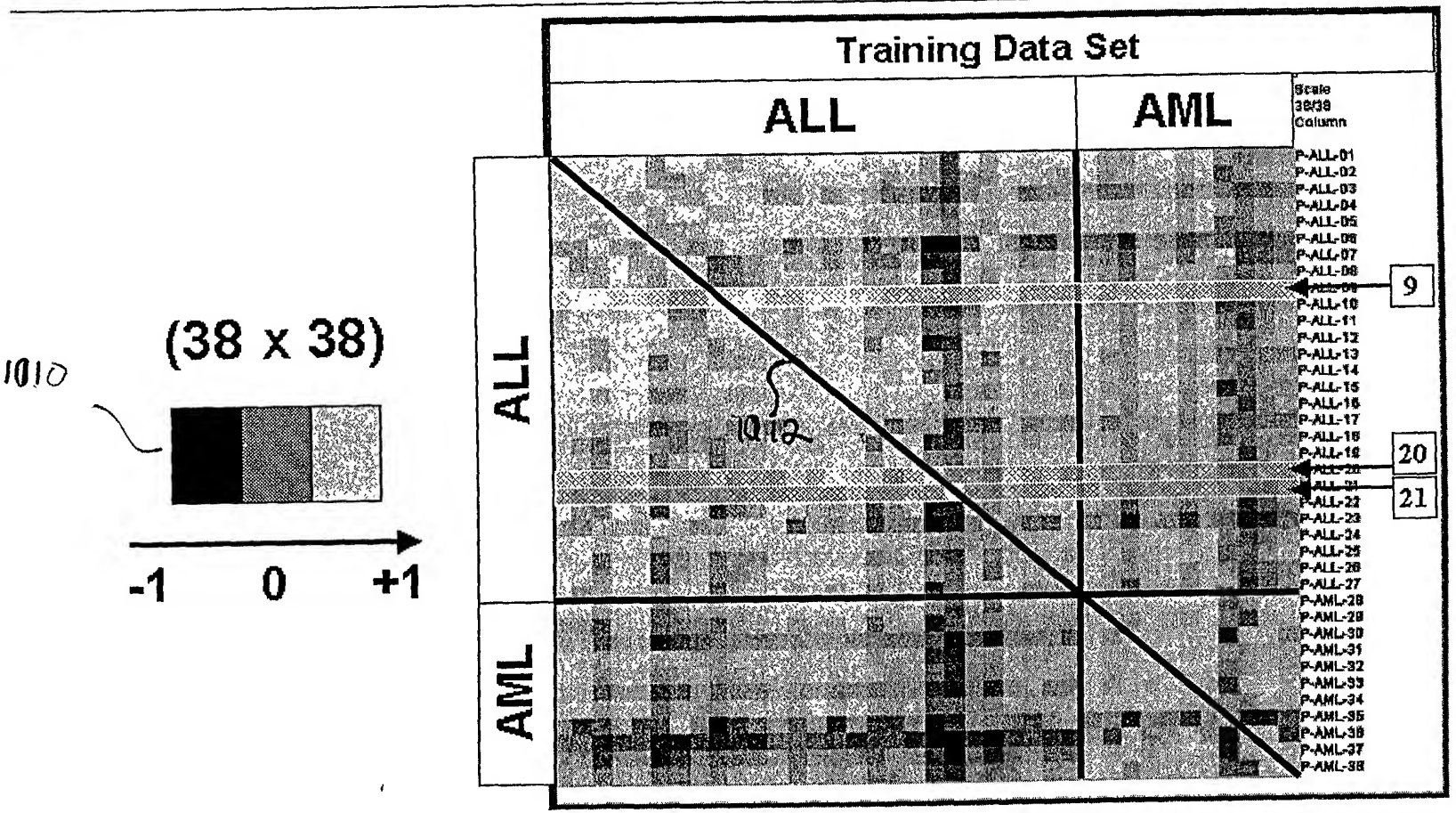


Figure 10B

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 23 of 79

20STEP169Z00T

Pearson Correlation Coefficients
(Affymetrix Comparative Data)

1003

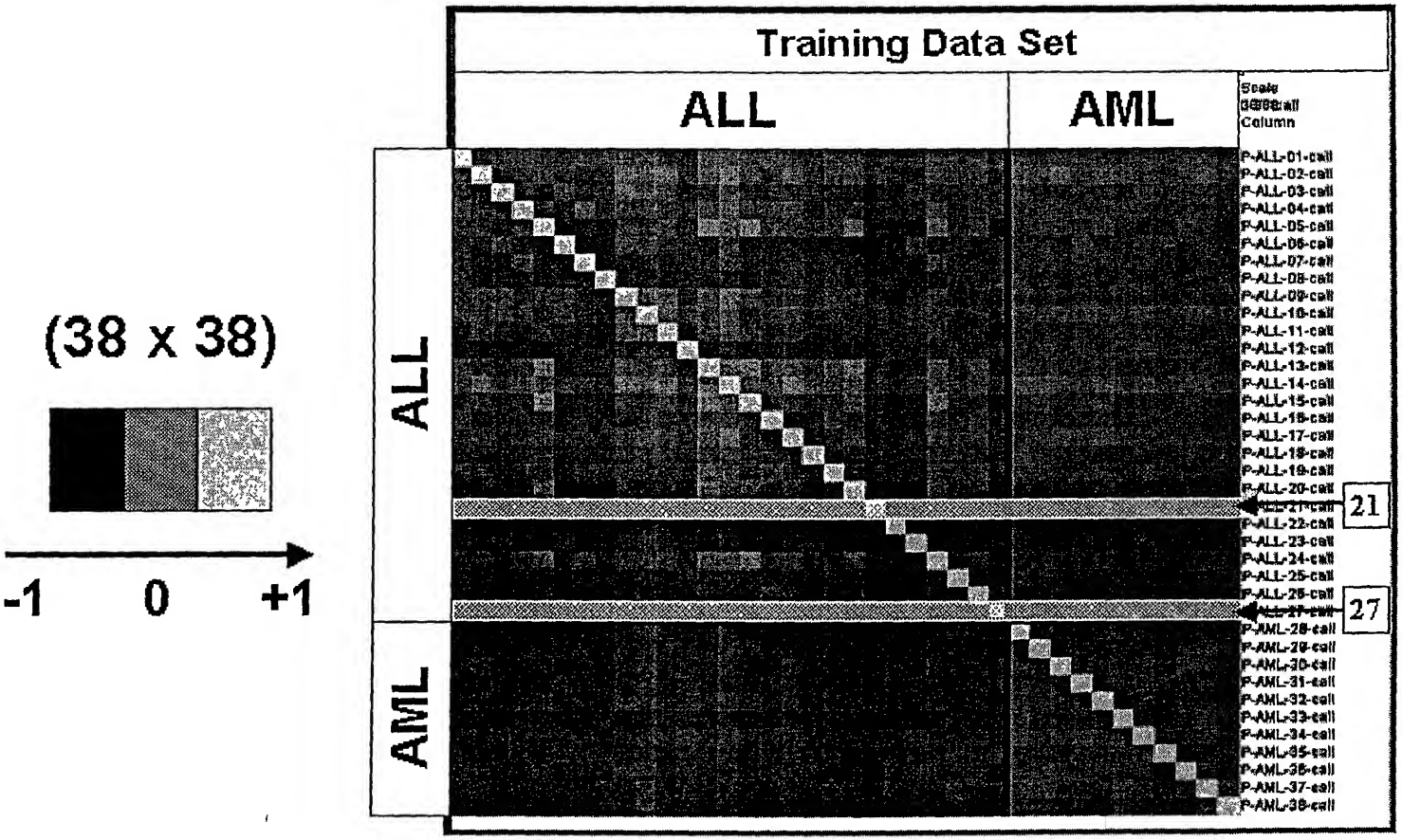


Figure 10C

2218917

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 24 of 79

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 25 of 79

1100

Final Result

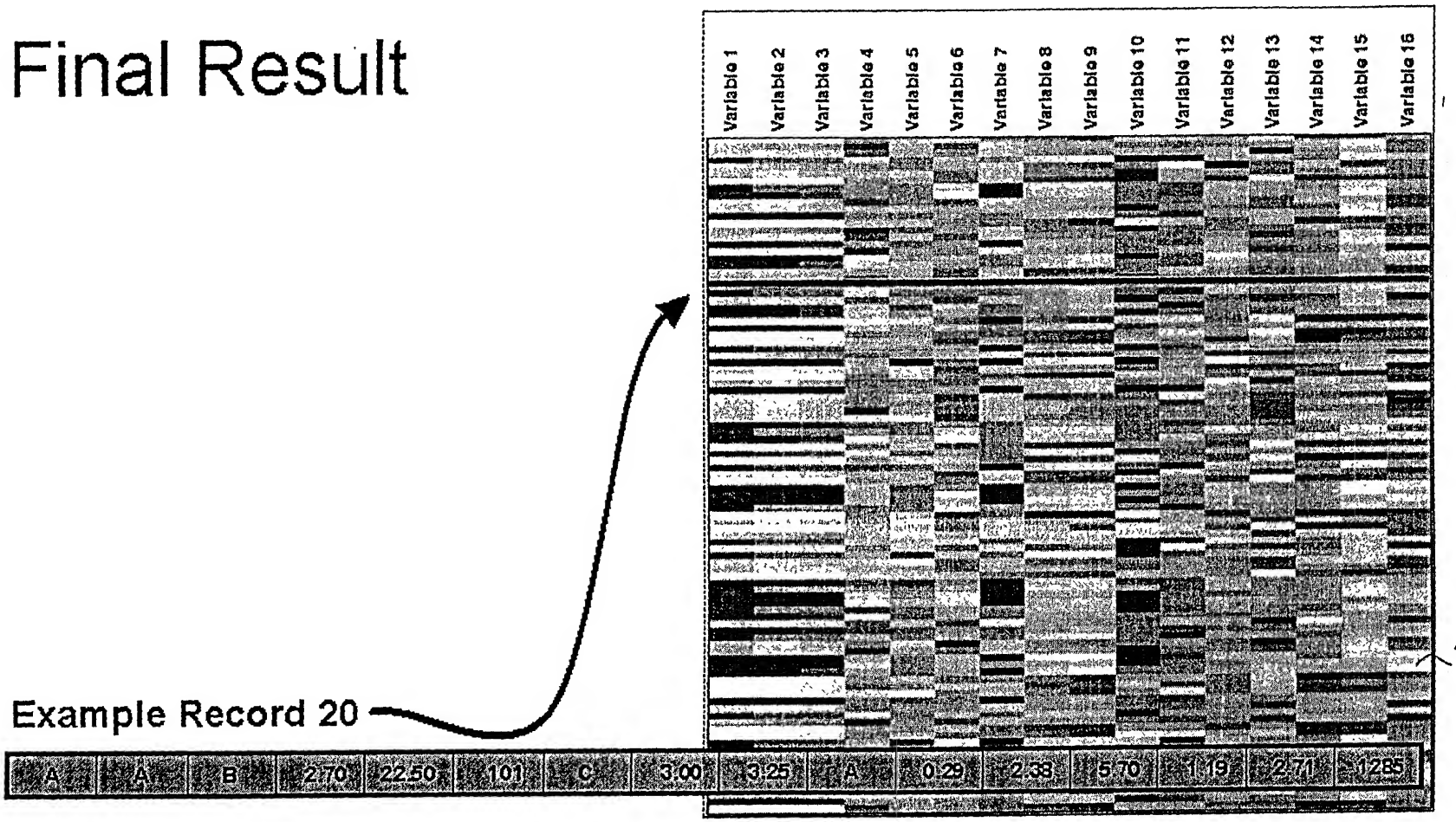


Figure 11A

20s t20 * t692t00t

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 26 of 79

Sort

– Variable 1

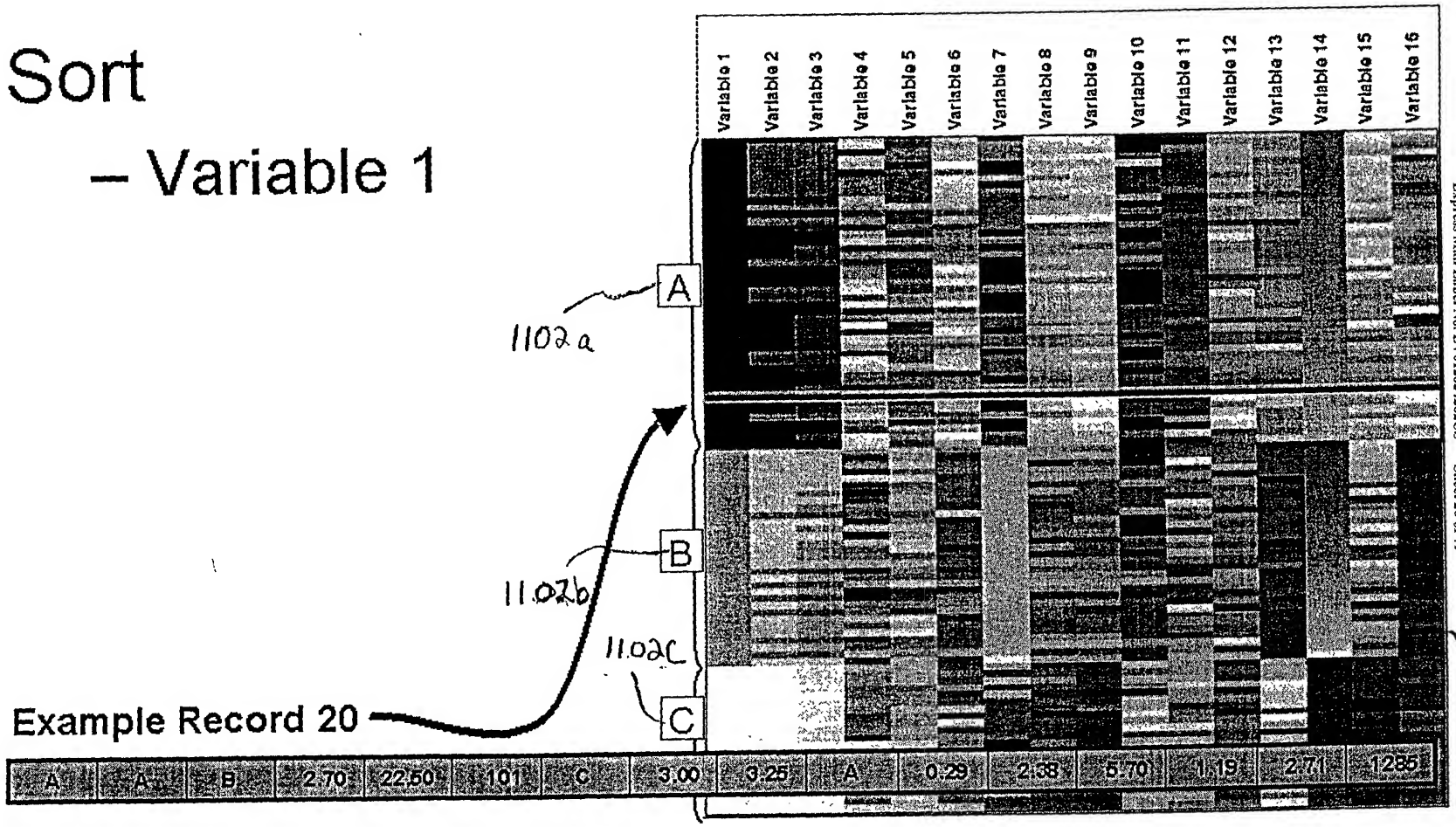


Figure 11B

2218917

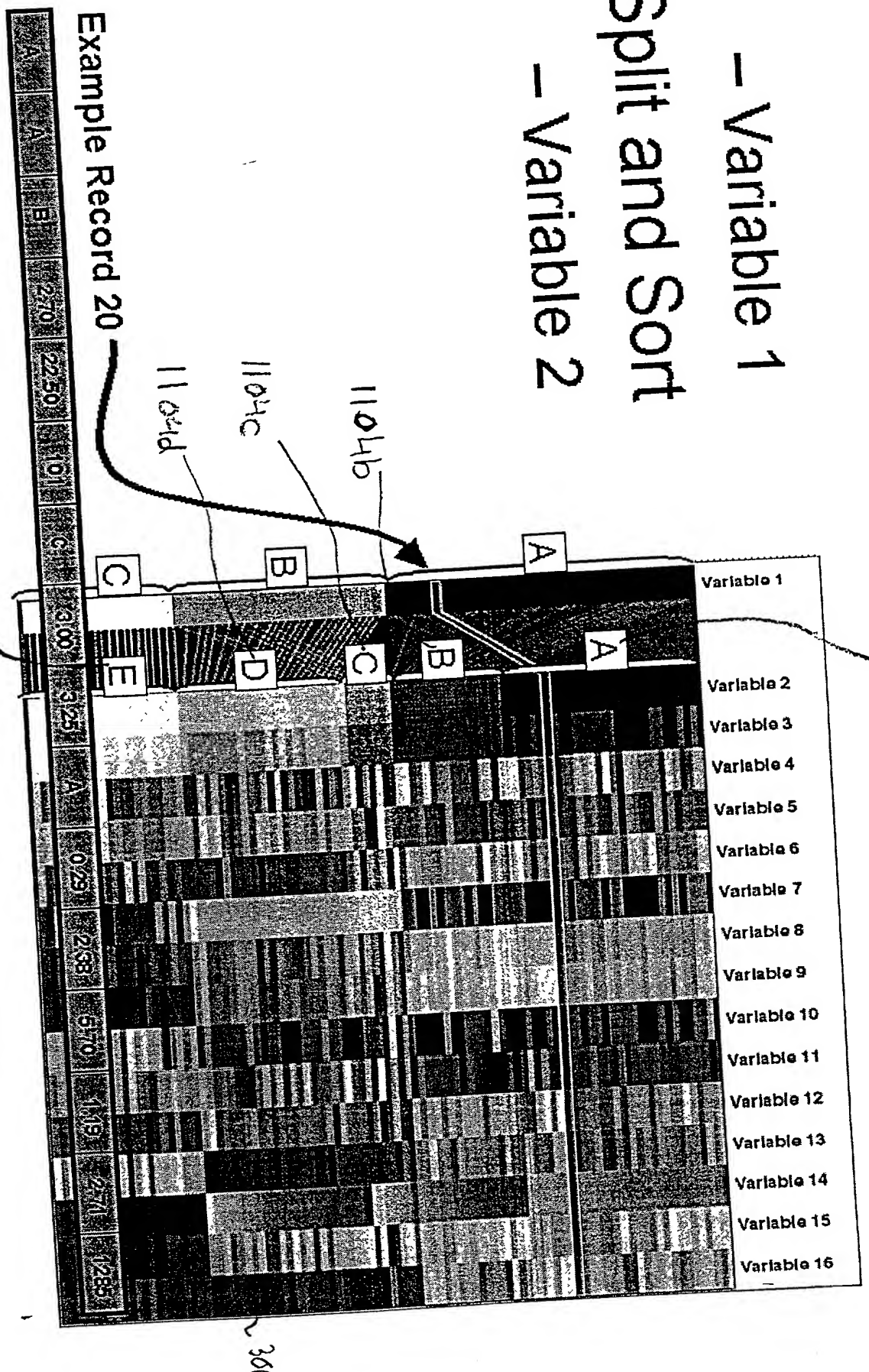


Figure 11C

11/8/2017

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 28 of 79

– Variable 1
– Variable 2
Split and Sort
– Variable 3

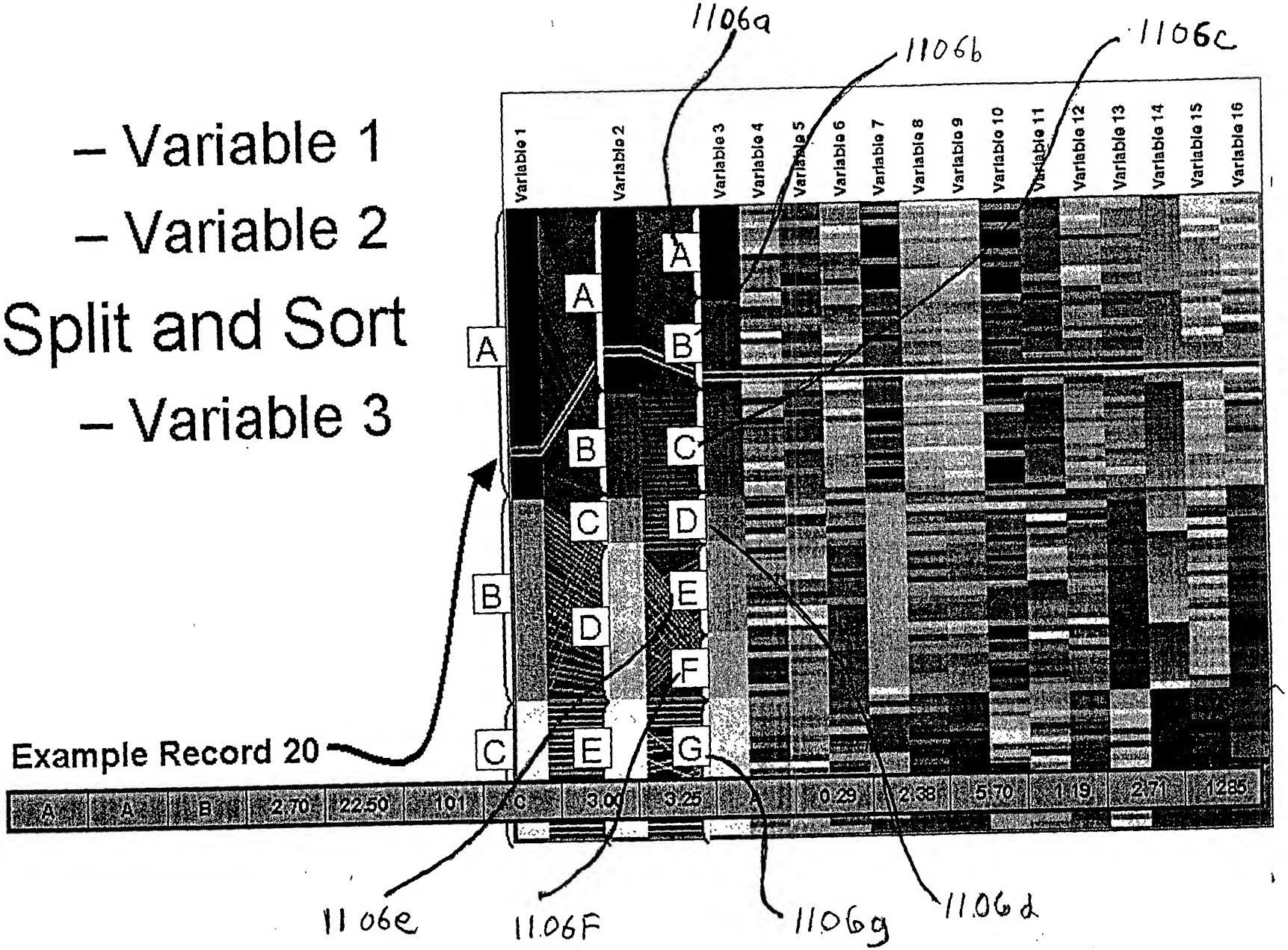
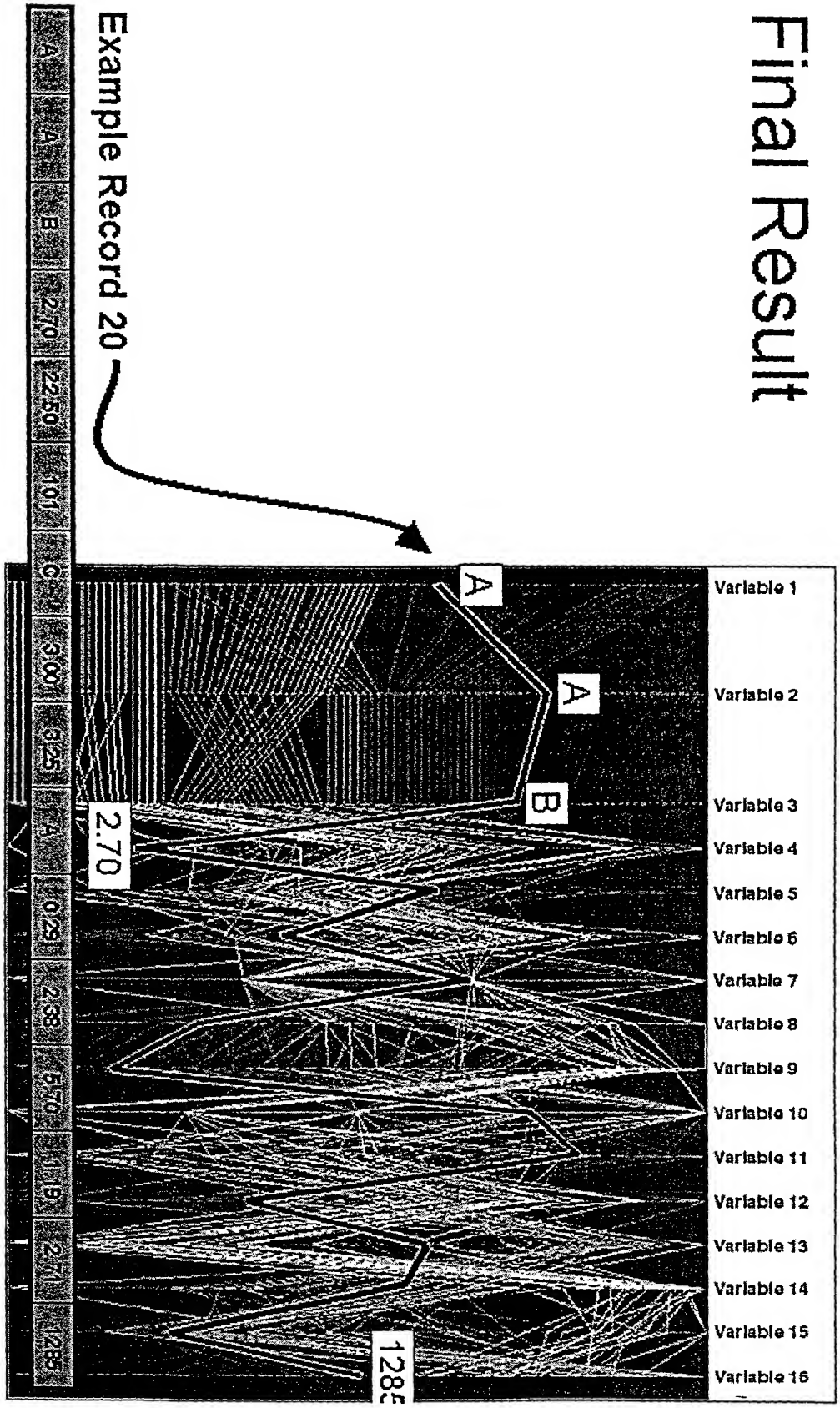


Figure 11D


Final Result
Example Record 20
A
A
B
2.70
1285
Variable 1
Variable 2
Variable 3
Variable 4
Variable 5
Variable 6
Variable 7
Variable 8
Variable 9
Variable 10
Variable 11
Variable 12
Variable 13
Variable 14
Variable 15
Variable 16

2218917

20SF20 "h697200F

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 30 of 79

Example record showing the underlying connections

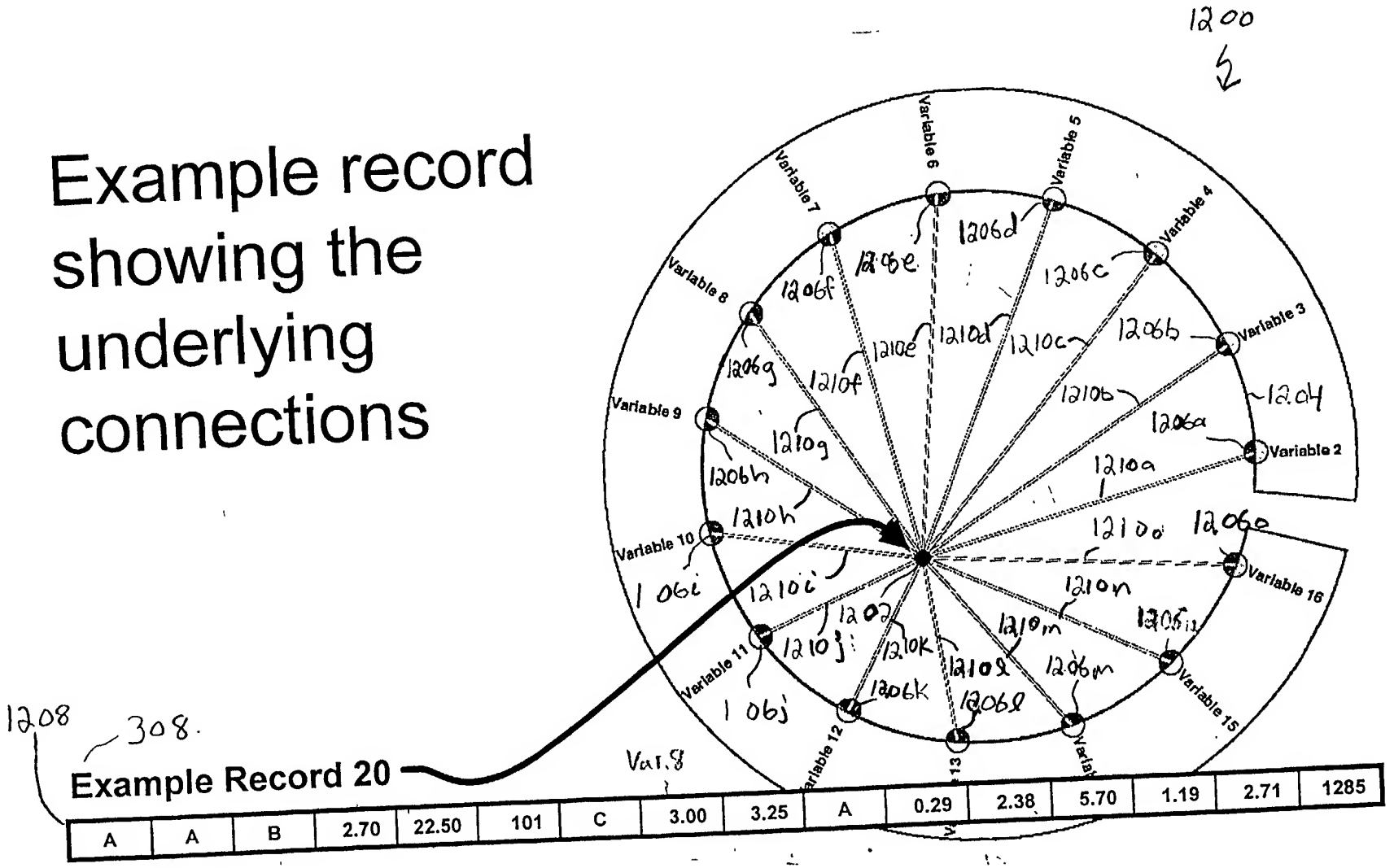


Figure 12A

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 31 of 79

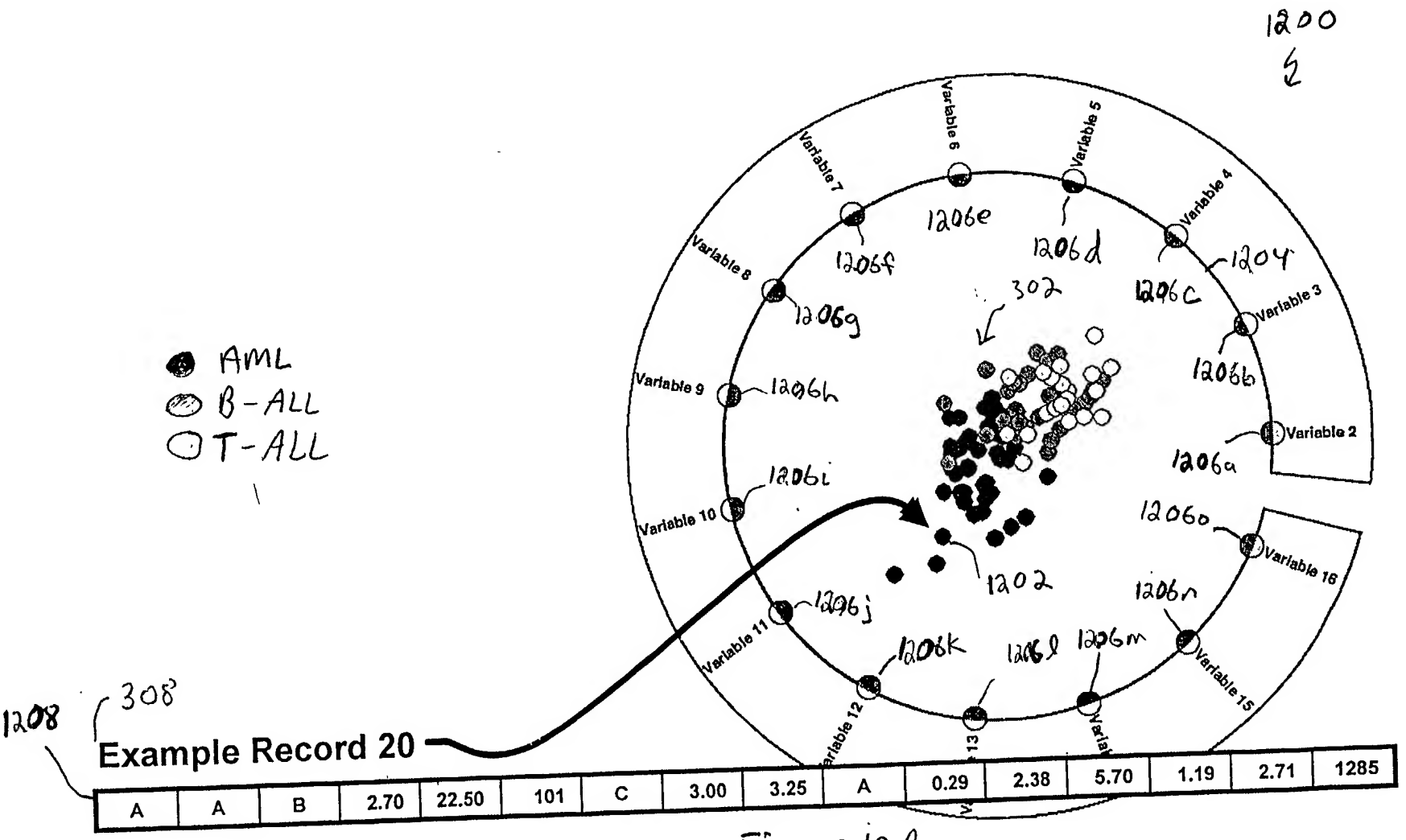


Figure 12B

2259888

20ST20 "h692400F

Visual Classifier

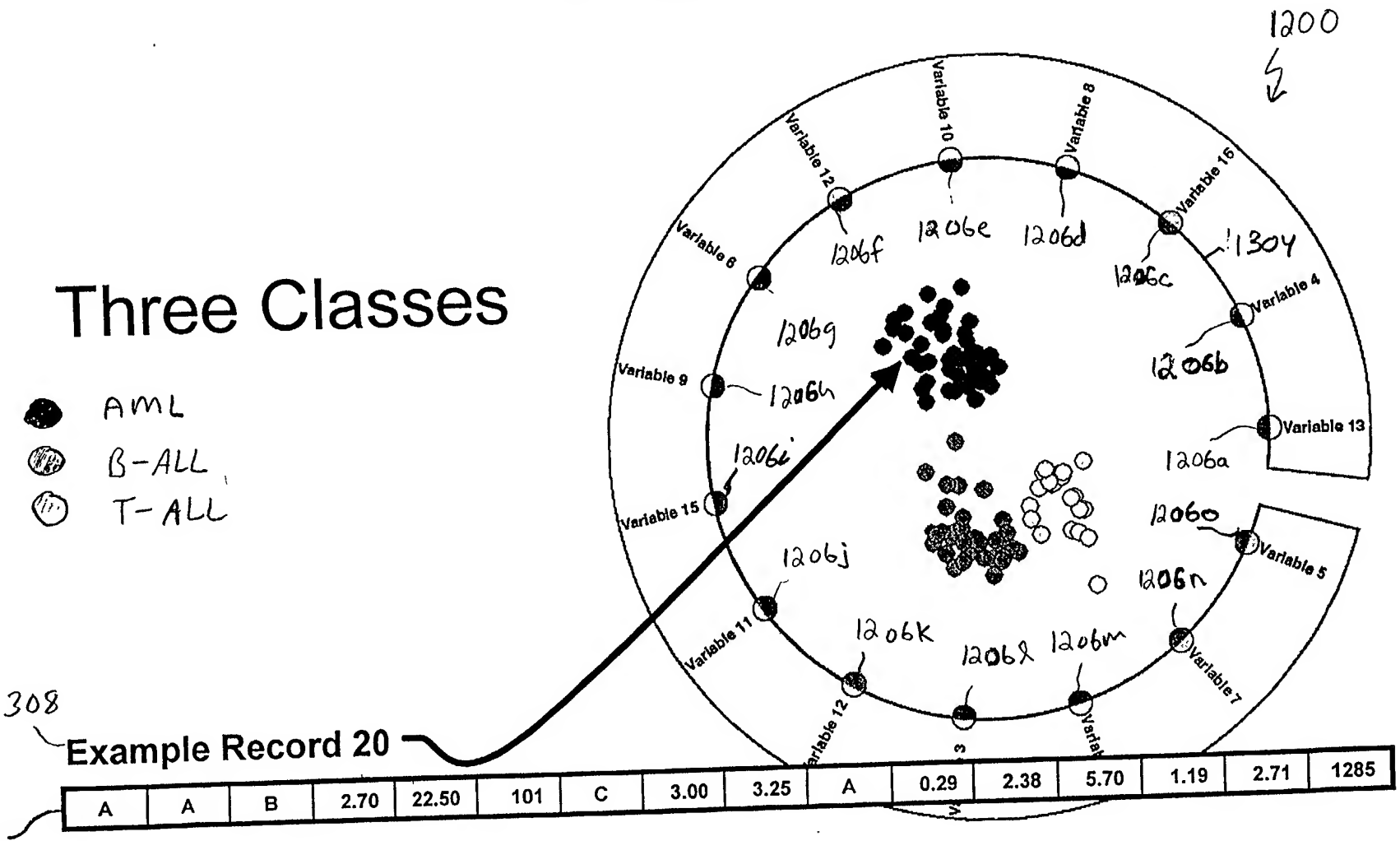


Figure 12C

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 32 of 79

2259888

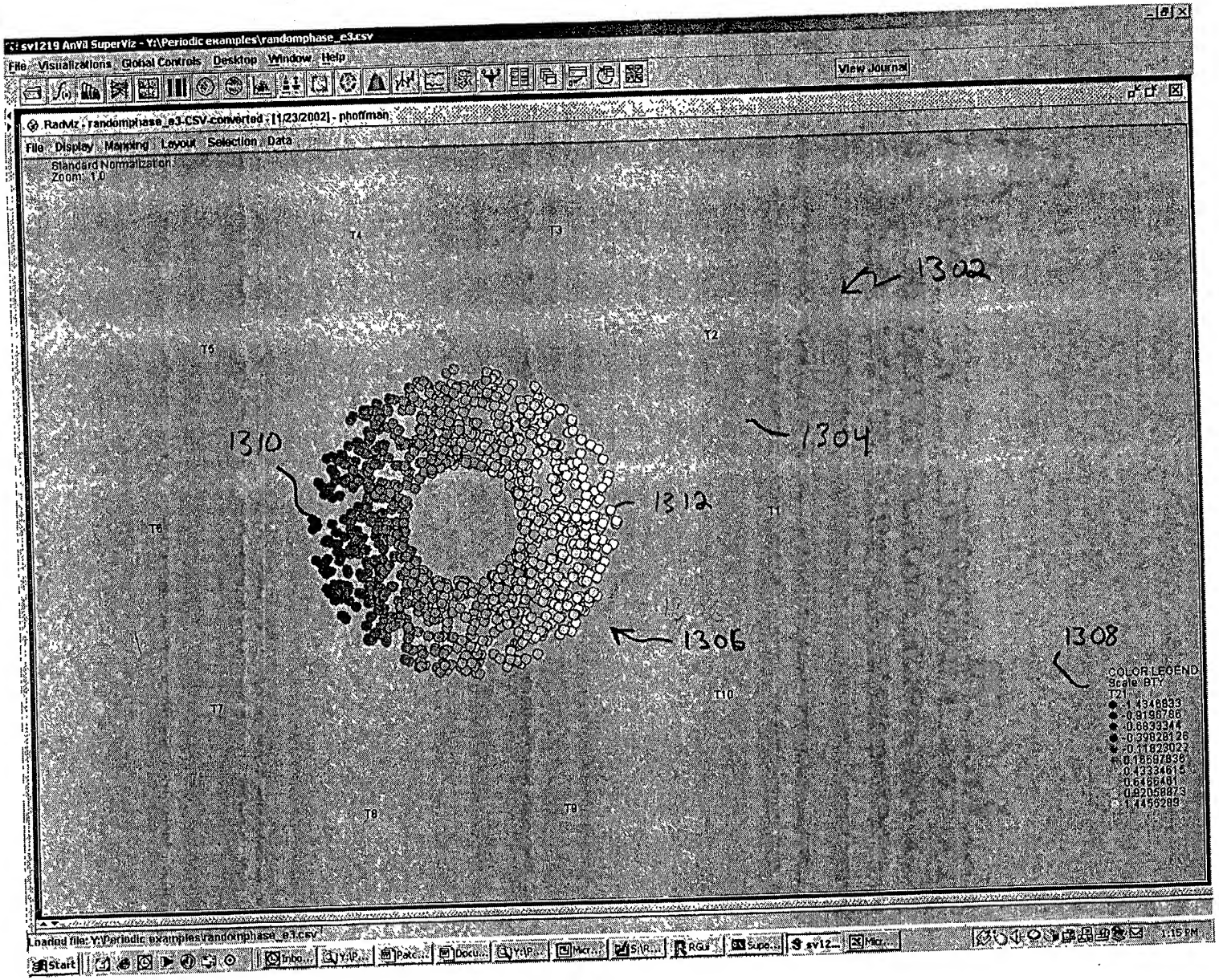

sv1219 AnVil SuperViz - Y:\Periodic examples\randomphase_e3.csv
File Visualizations Global Controls Desktop Window Help
View Journal
RadViz - randomphase_e3.CSV.converted - [1/23/2002] - phoffman
File Display Mapping Layout Selection Data
Standard Normalization
Zoom: 1.0
T1
T2
T3
T4
T5
T6
T7
T8
T9
T10
1302
1304
1306
1308
1310
1312
COLOR LEGEND
Scale: BTY
T21
1.4346833
0.9196786
0.6833344
0.3982812 6
0.11823022
-0.18697836
-0.4333461 5
-0.6466401
-0.8205987 3
-1.4456289
Loading file: Y:\Periodic examples\randomphase_e3.csv
Start
1:15 PM

Figure 13A

20STP20 "h692007T

1301

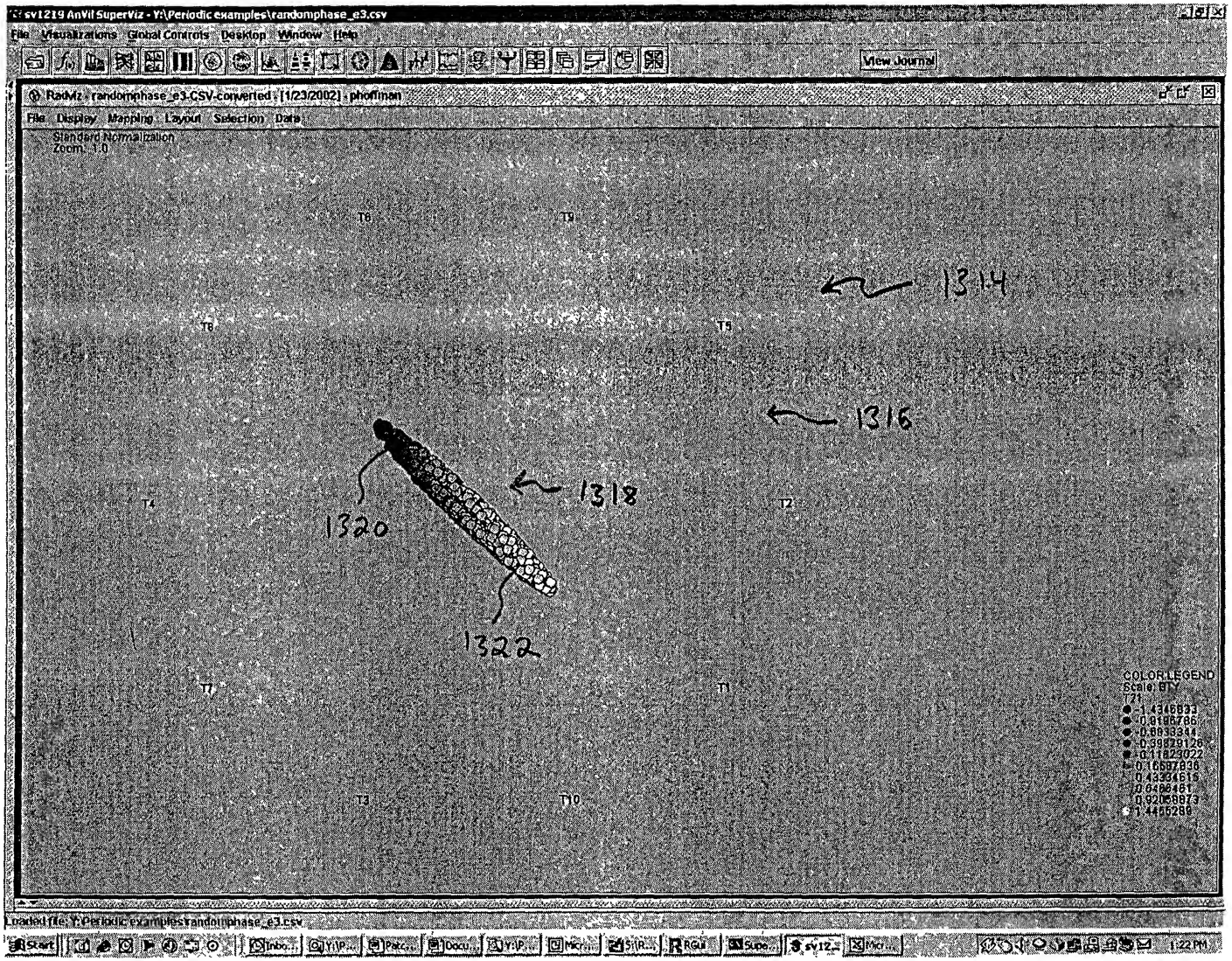


Figure 13B

2269307

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 34 of 79

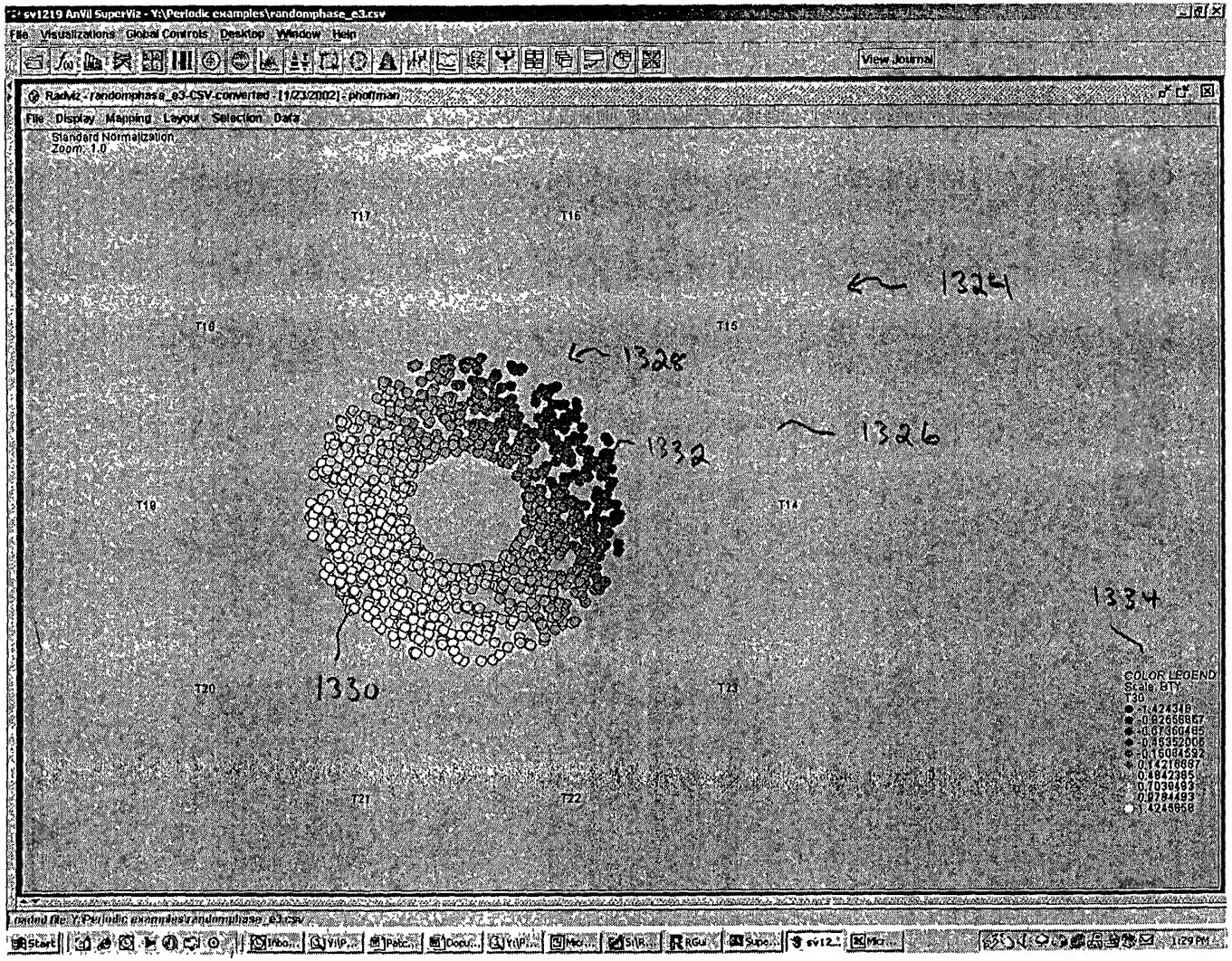


Figure 13C

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 35 of 79

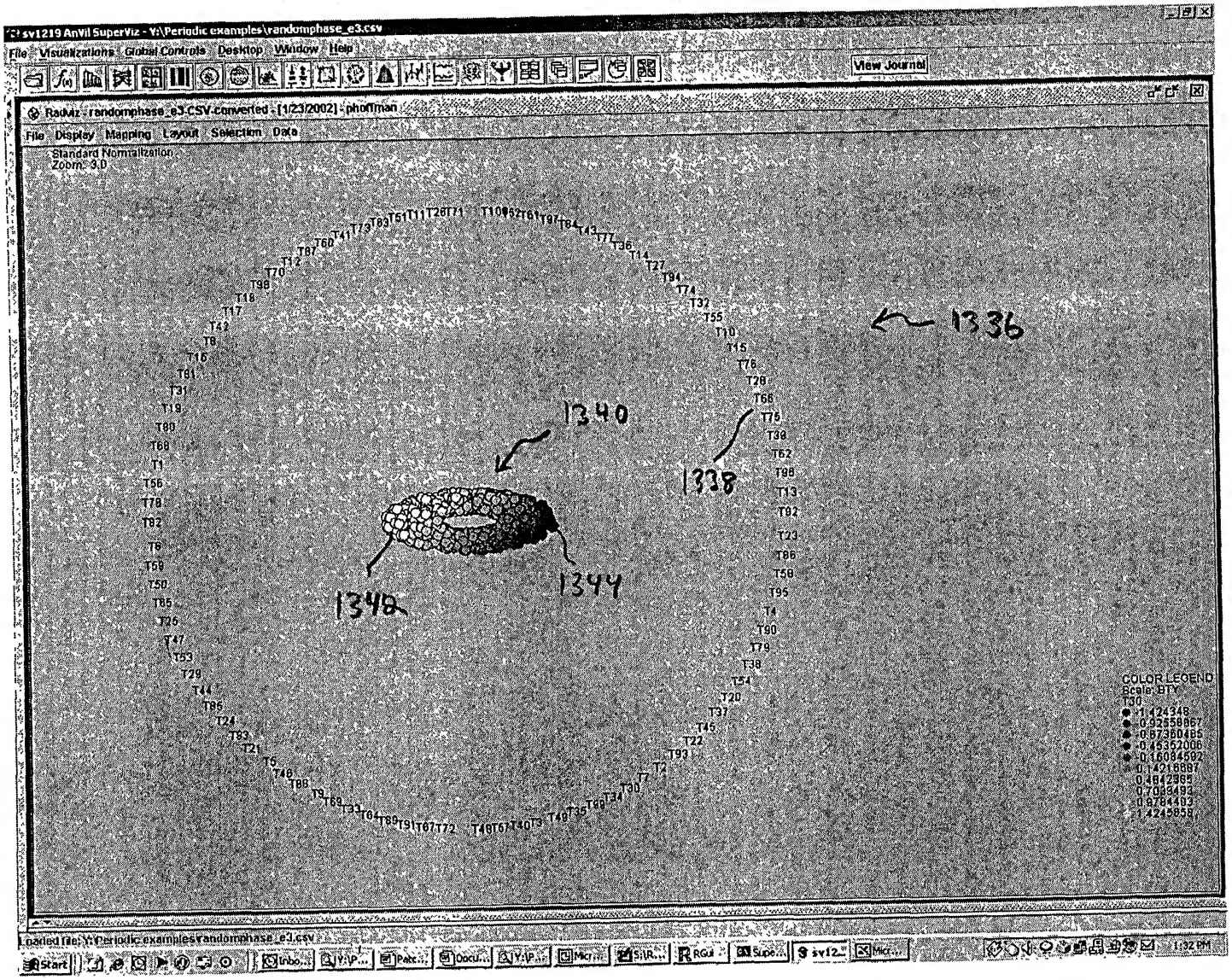


Figure 13D

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 36 of 79

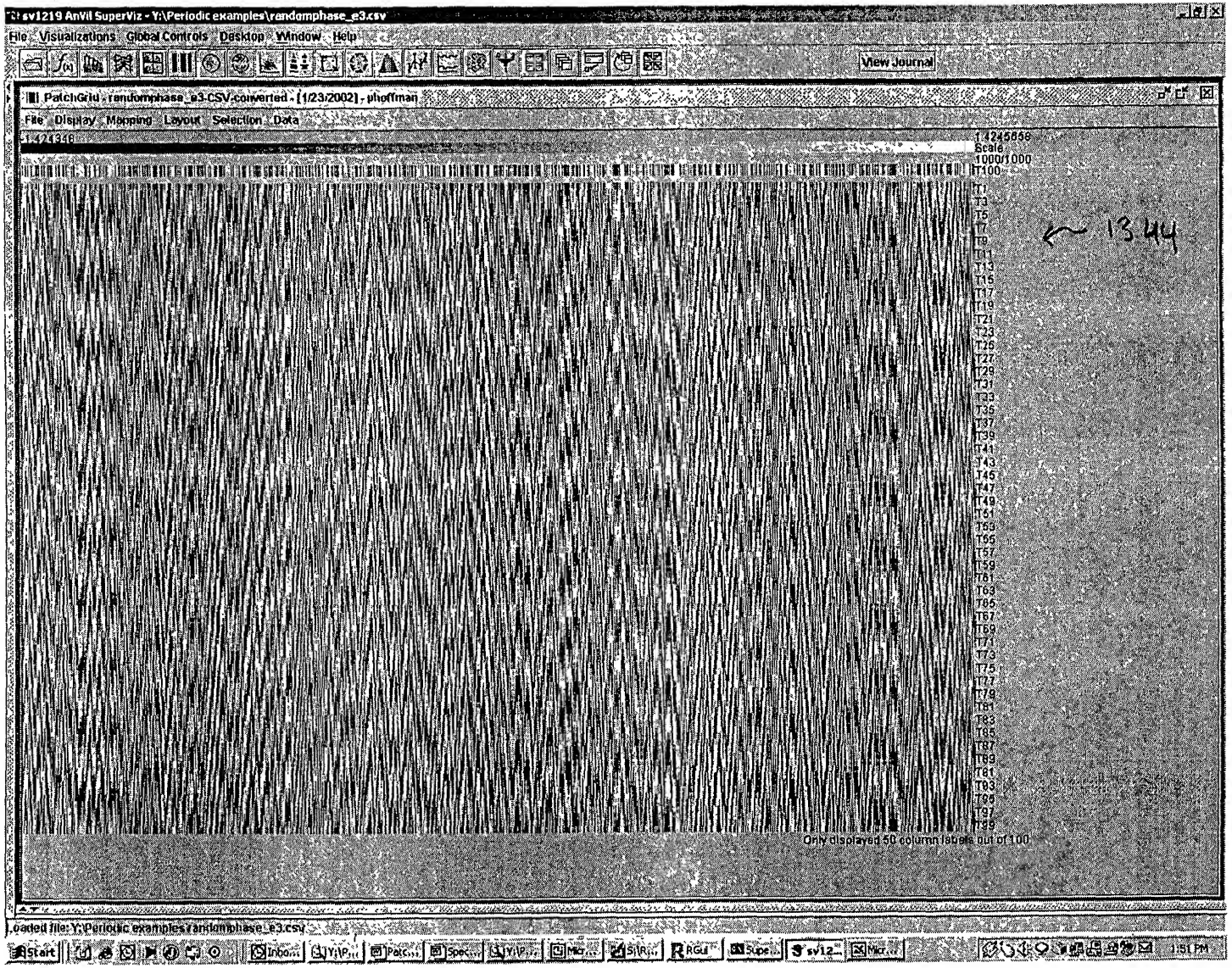

sv1219 AnVil SuperViz - Y:\Periodic examples\randomphase_e3.csv
File Visualizations Global Controls Desktop Window Help
View Journal
PatchGrid - randomphase_e3.CSV-converted - (1/23/2002) - phoffman
File Display Mapping Layout Selection Data
1.421348
1.4245656
Scale
1000/1000
T100
Only displayed 50 column labels out of 100
Loaded file: Y:\Periodic examples\randomphase_e3.csv
Start
1:51 PM

Figure 13B

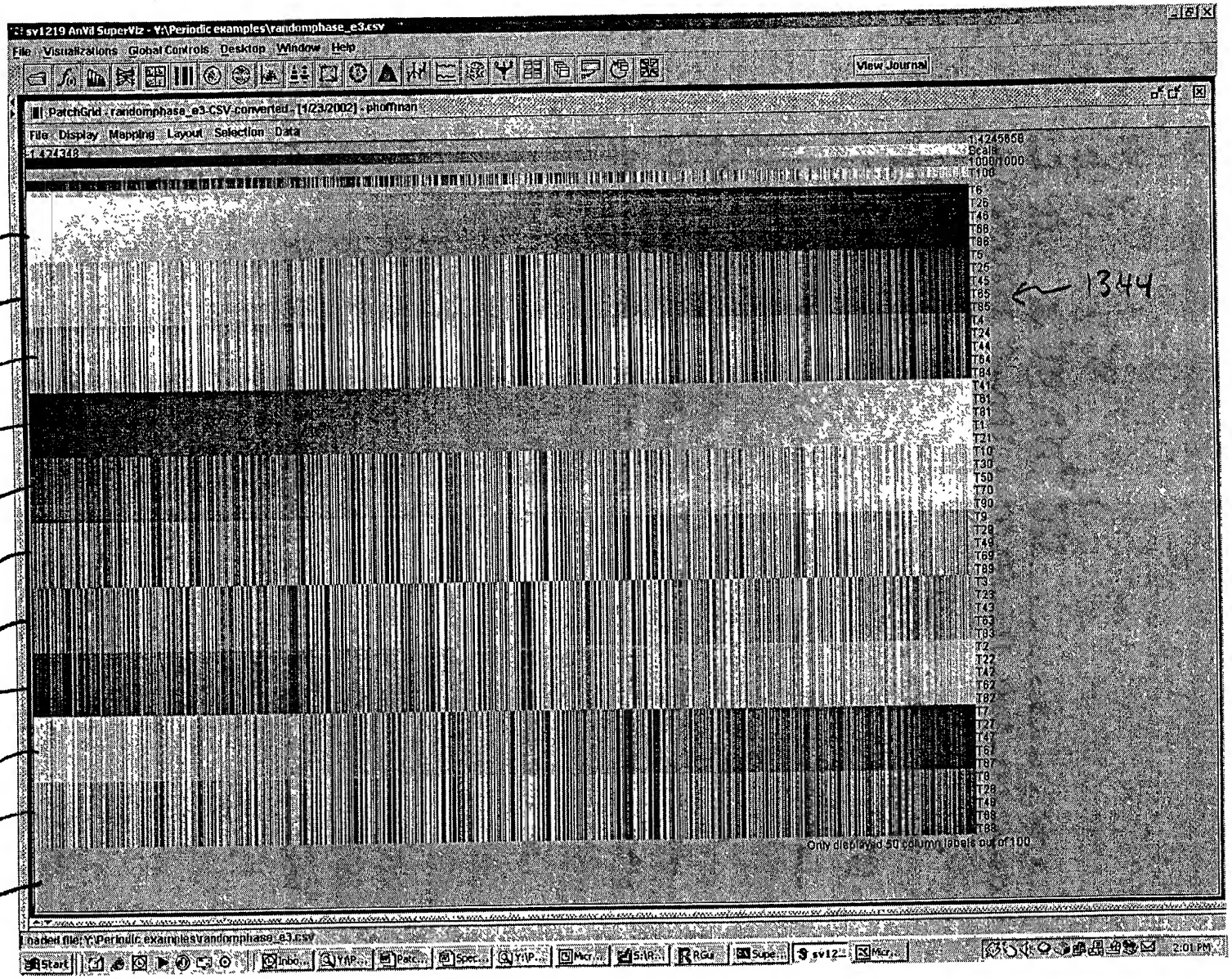


Figure 13F

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 38 of 79

20STP20 "h692c00T

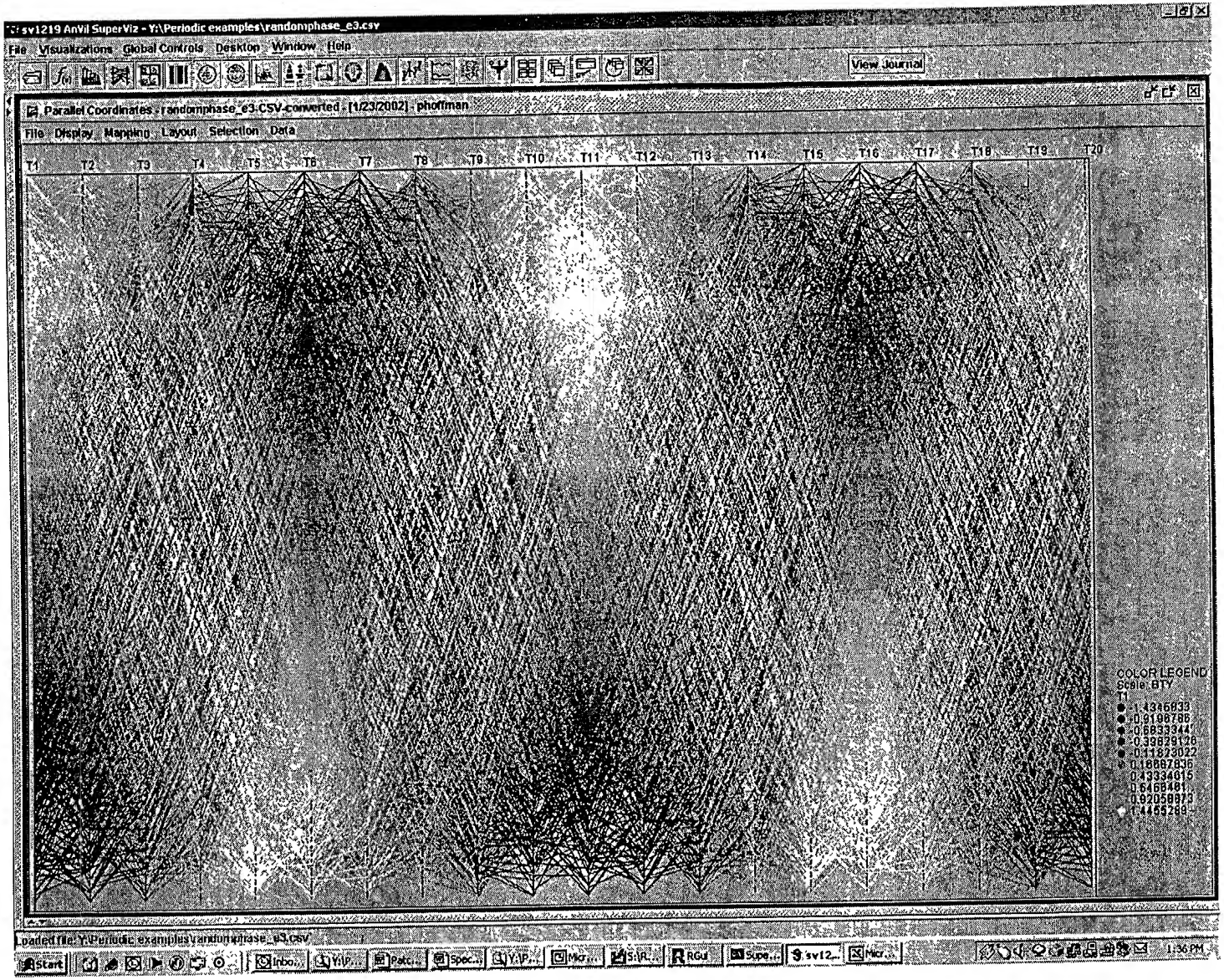


Figure 13G

20020-1692007

Transforming a rectangular to a radial layout

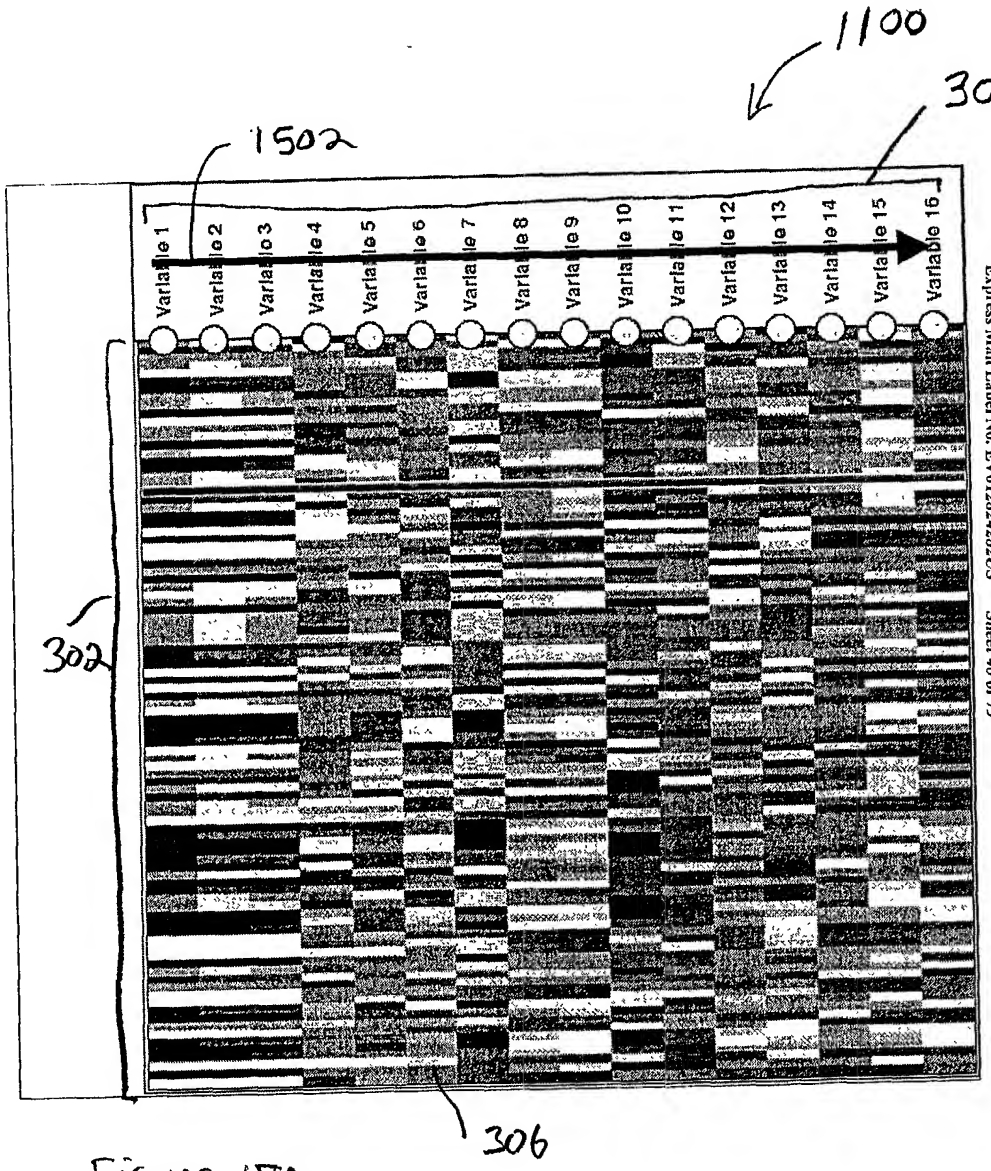


Figure 15A

D7598K8

20STEP "h892200F

Radial Layout

Transforming a rectangular to a radial layout

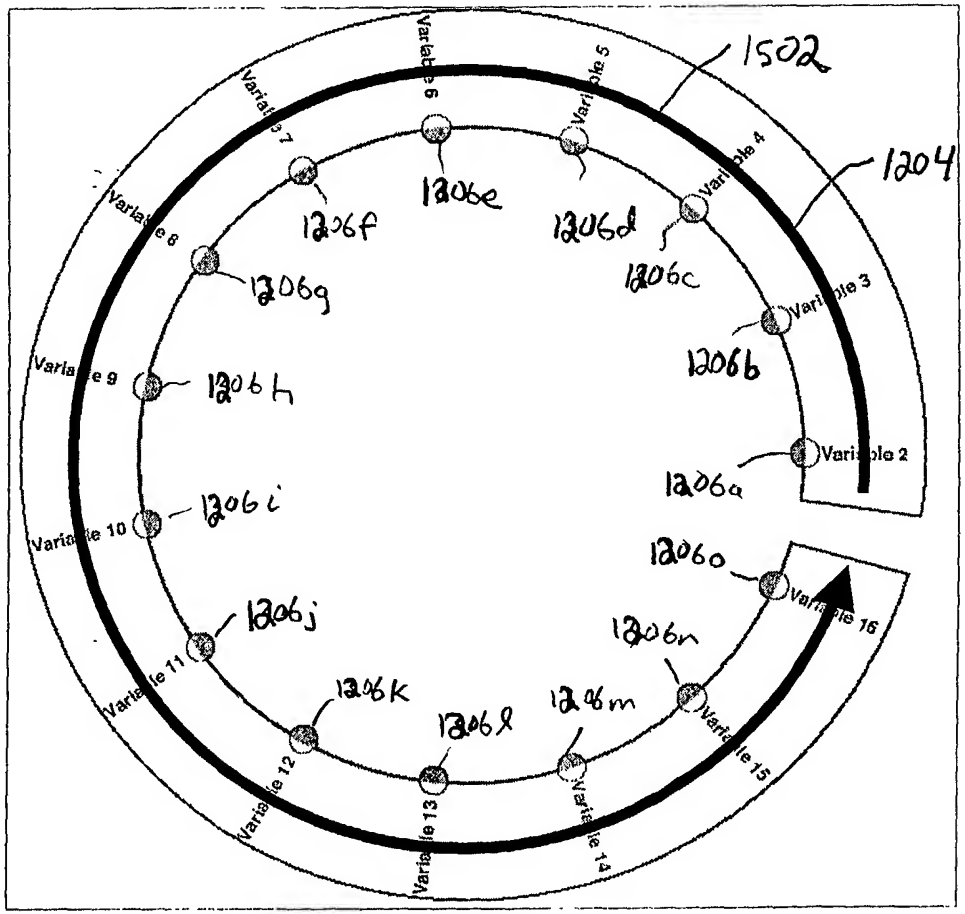


Figure 15B

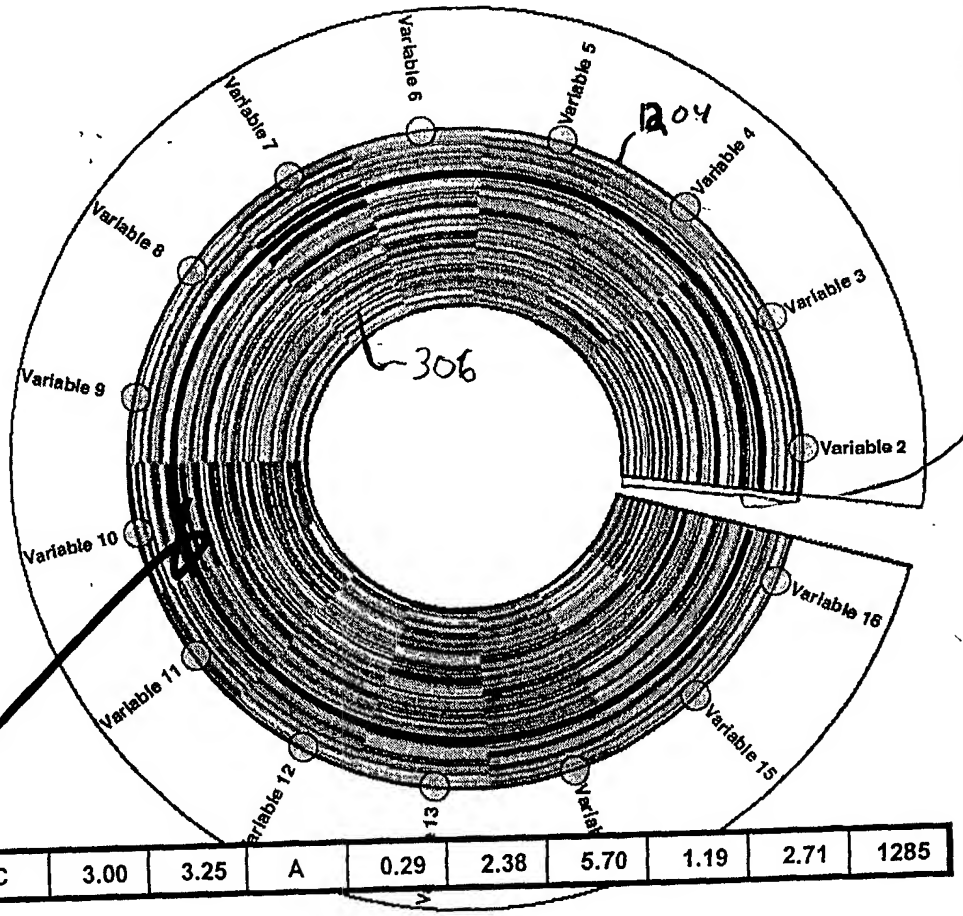
Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANY-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 41 of 79

22ea ocd

20520 "692007

Radial Layout

Transforming a rectangular to a radial layout



308

Example Record 20

A	A	B	2.70	22.50	101	C	3.00	3.25	A	0.29	2.38	5.70	1.19	2.71	1285

1208

Figure 15 C

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 42 of 79

Shape Control – Before Moving Anchors

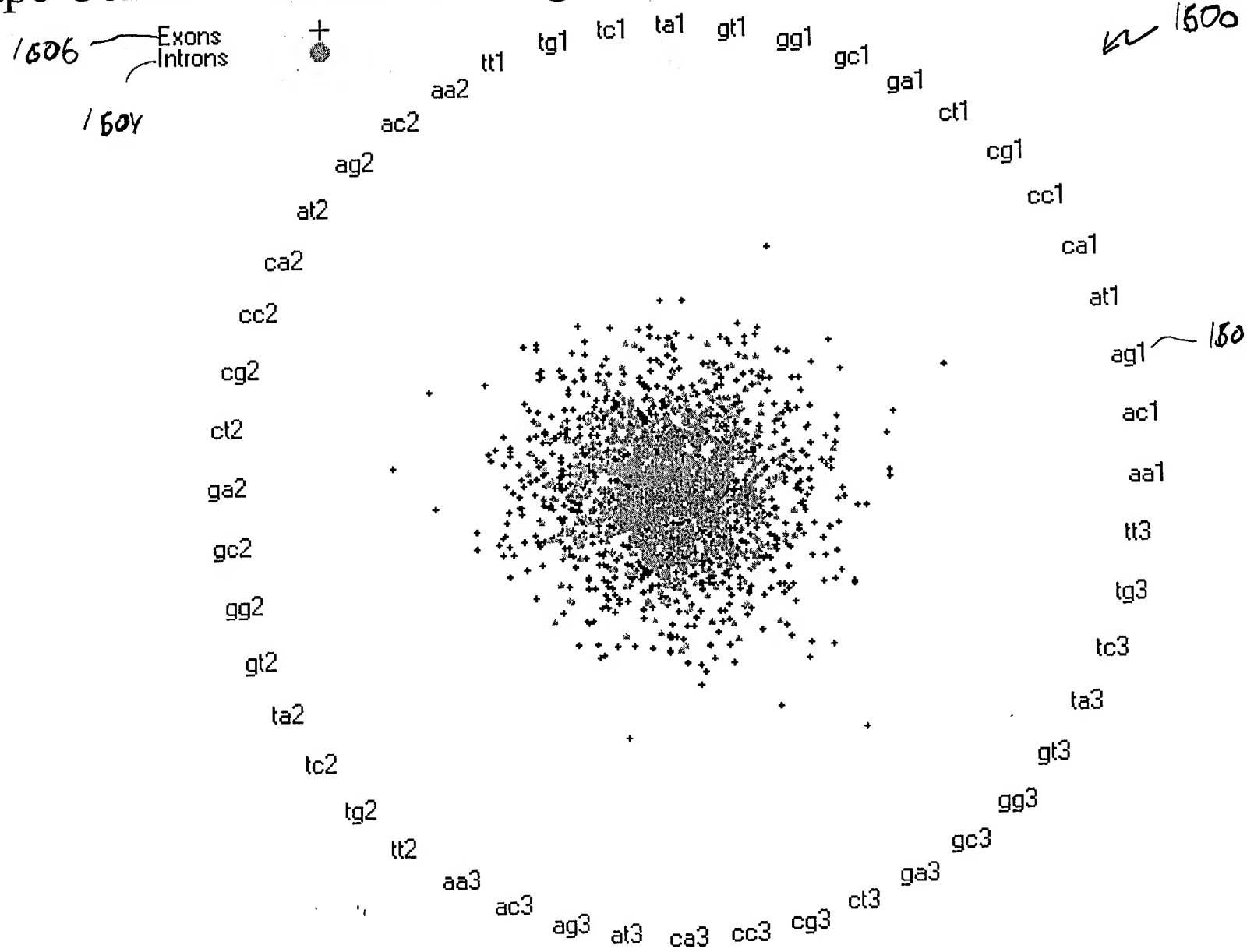


Figure 16A

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 43 of 79

Shape Control – Ellipse

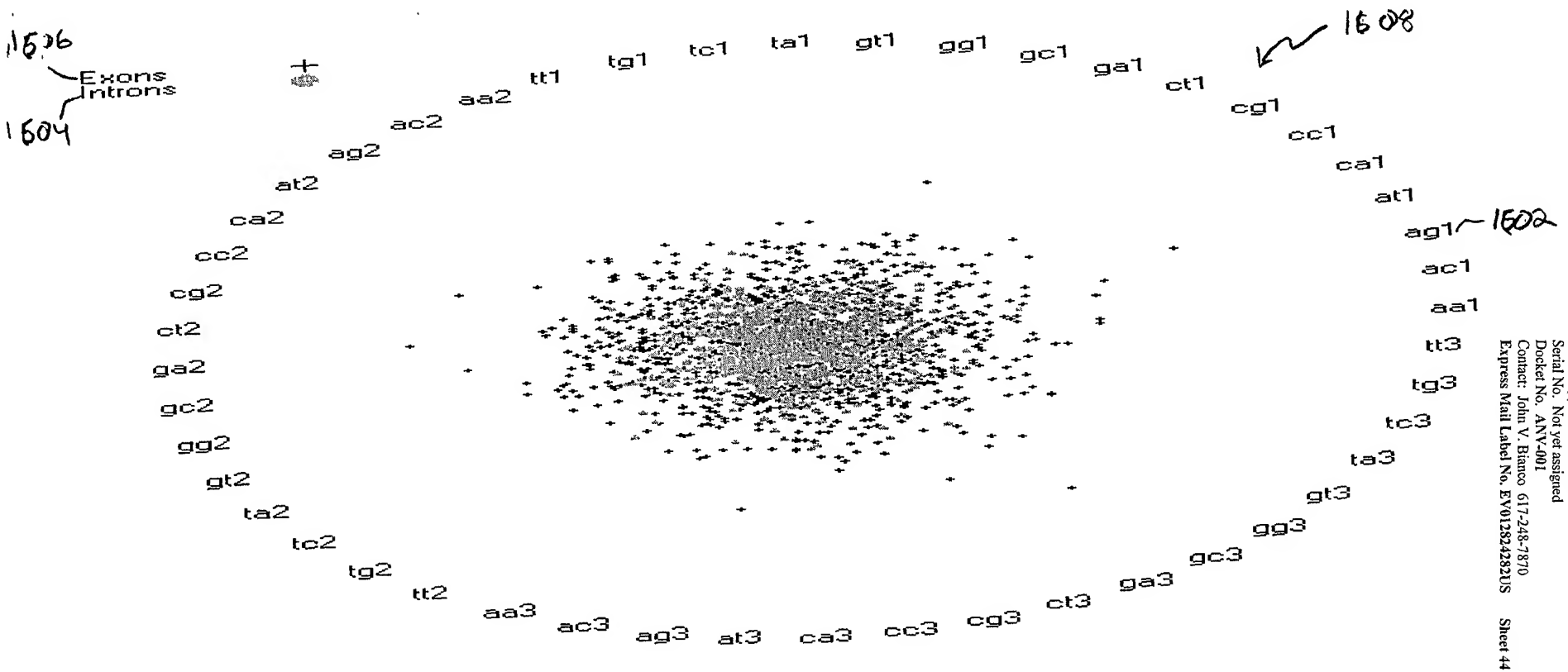


Figure 16B

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 44 of 79

2218917

Shape Control – Arbitrary Spline

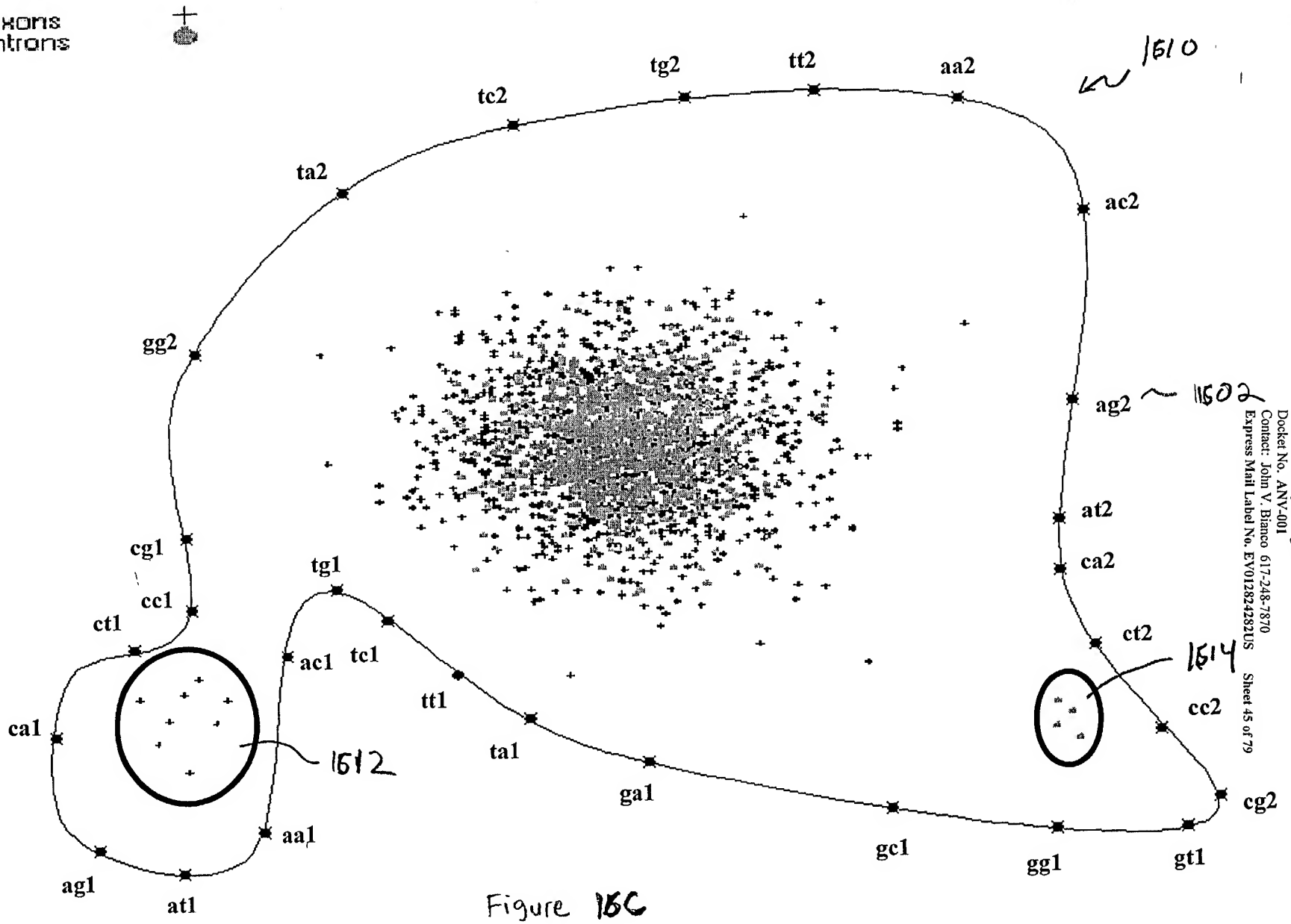


Figure 16C

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824292US Sheet 45 of 79

20520 "h69Z00T

1620

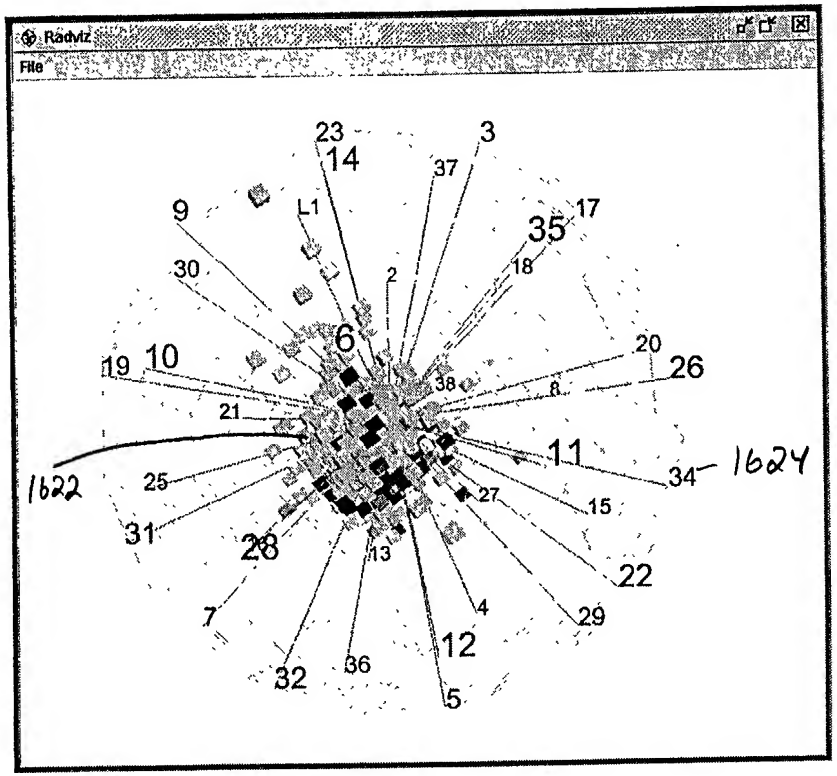


Fig. 16D

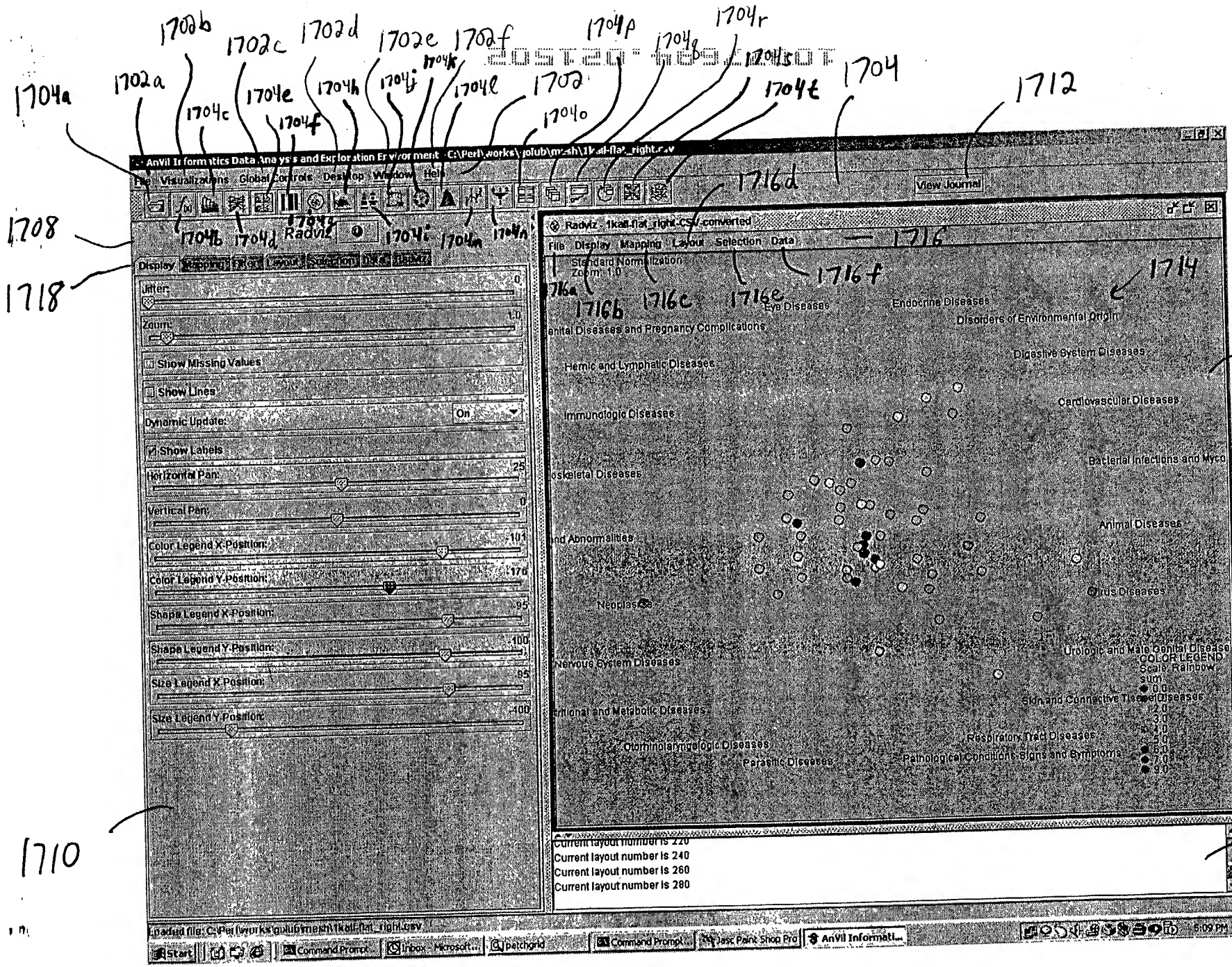


Figure 17

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 47 of 79

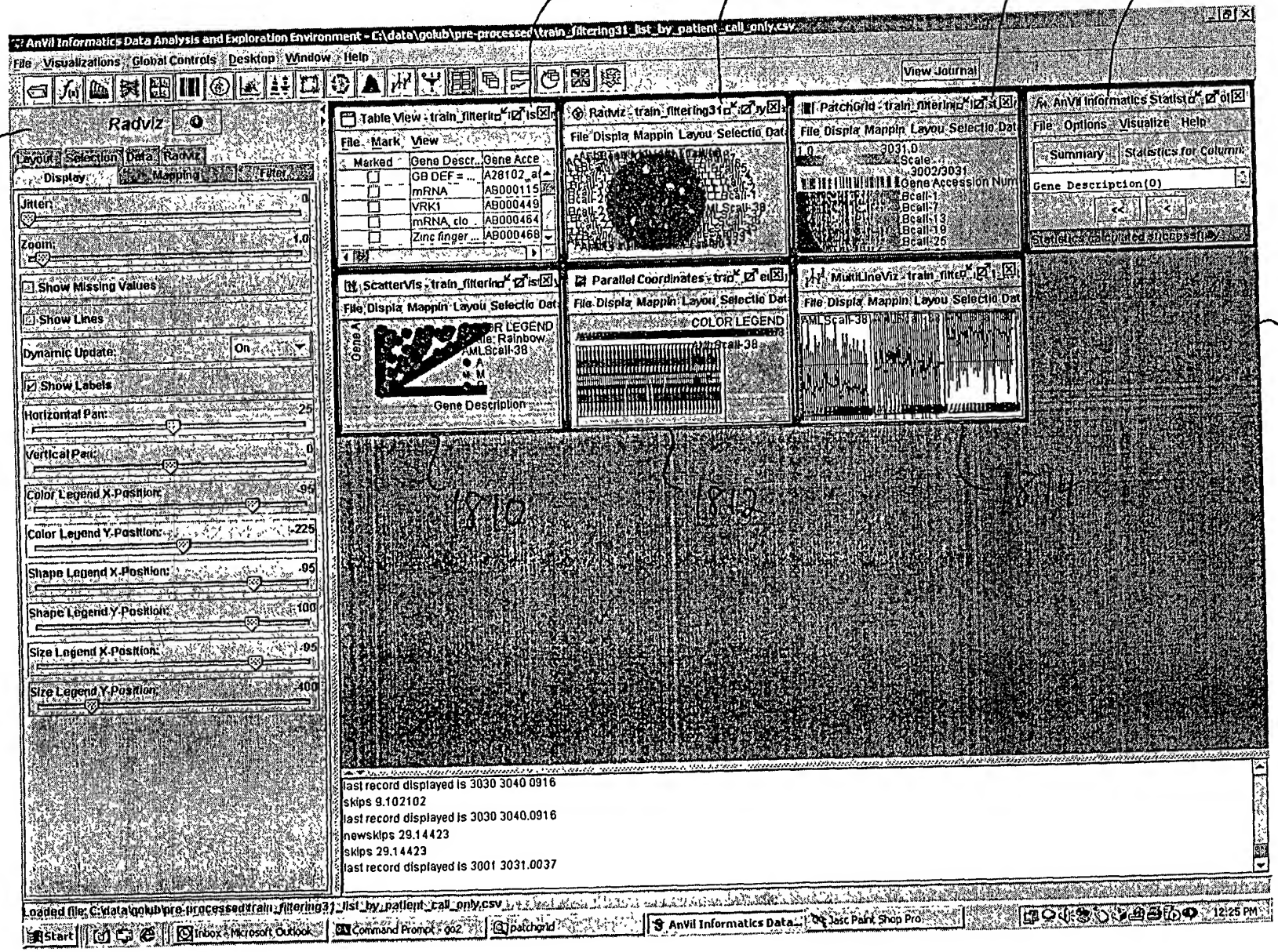


Figure 18

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 48 of 79

1900

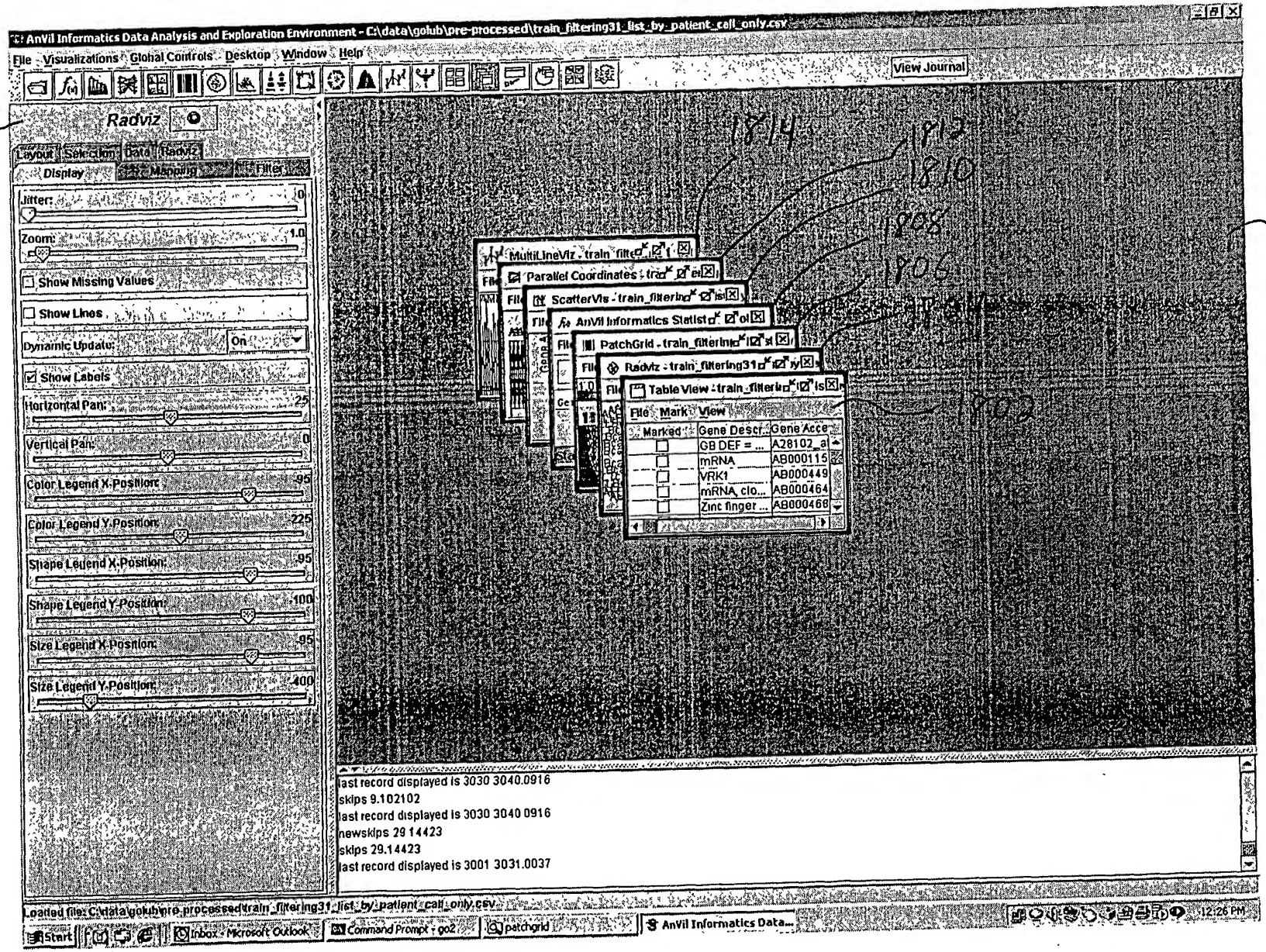


Figure 19

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 49 of 79

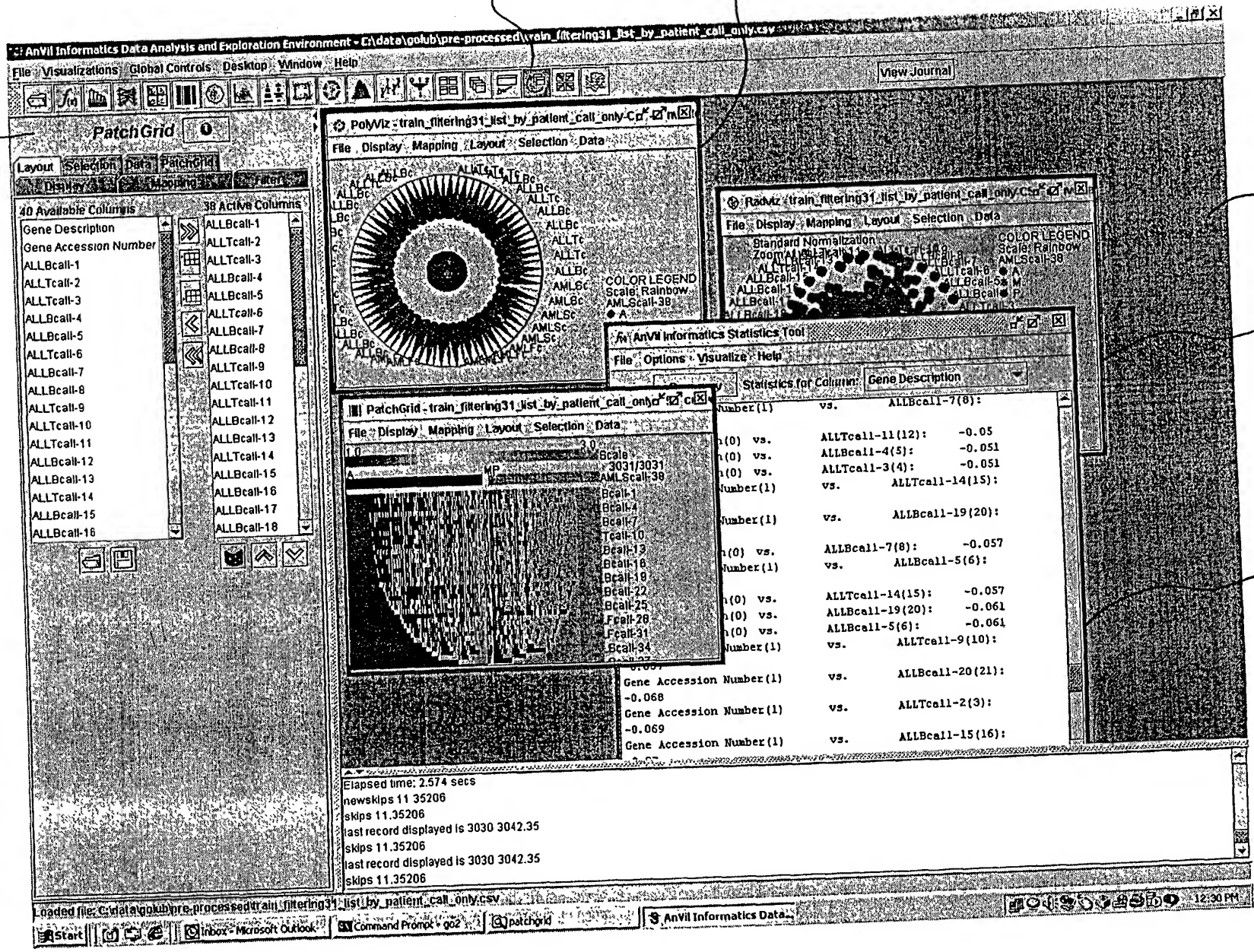


Figure 20

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No Not yet assigned
Docket No. ANV-001
Contact: John V Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 50 of 79

20160 h692007

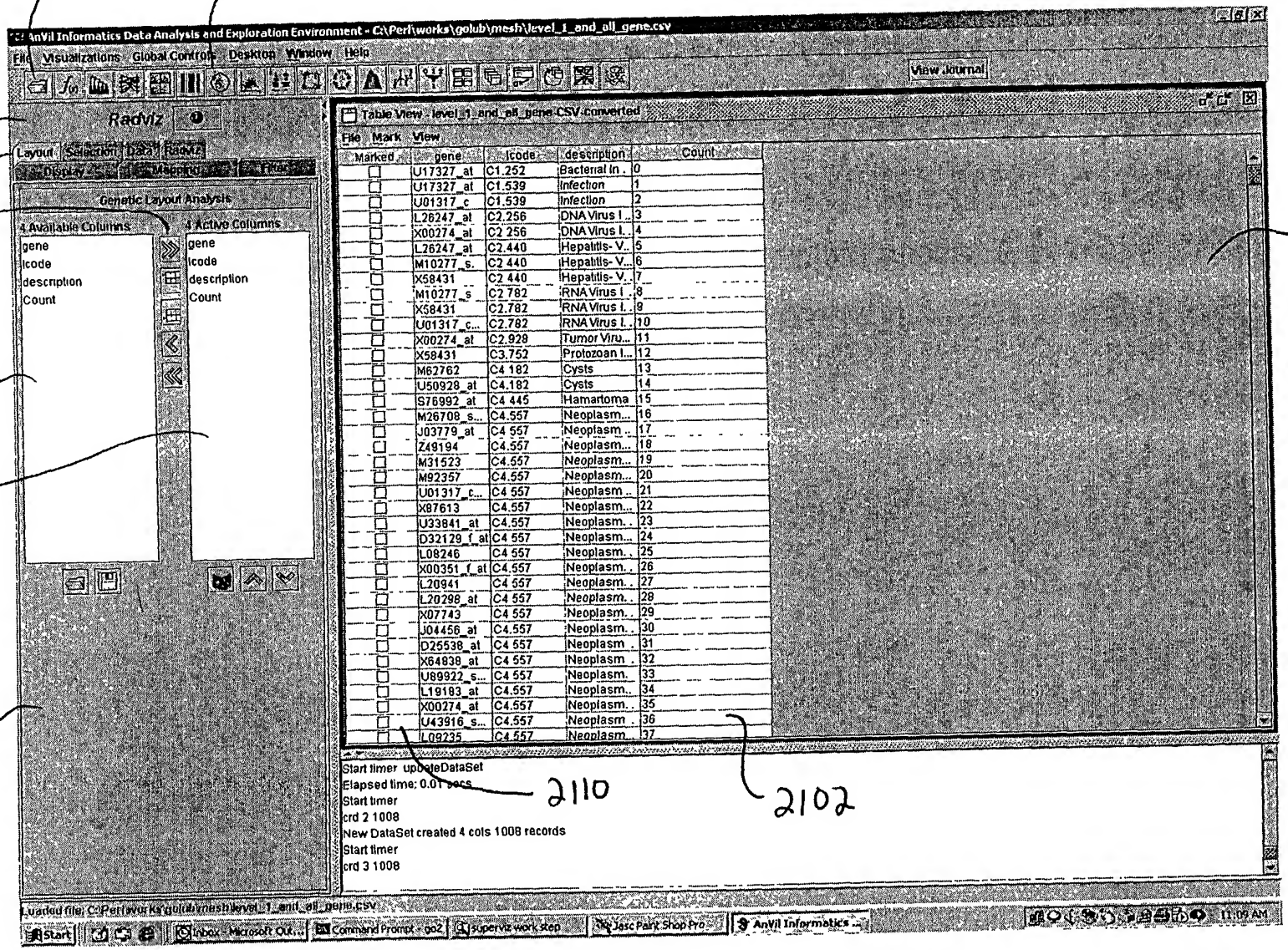


Figure 21

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 51 of 79

2174305

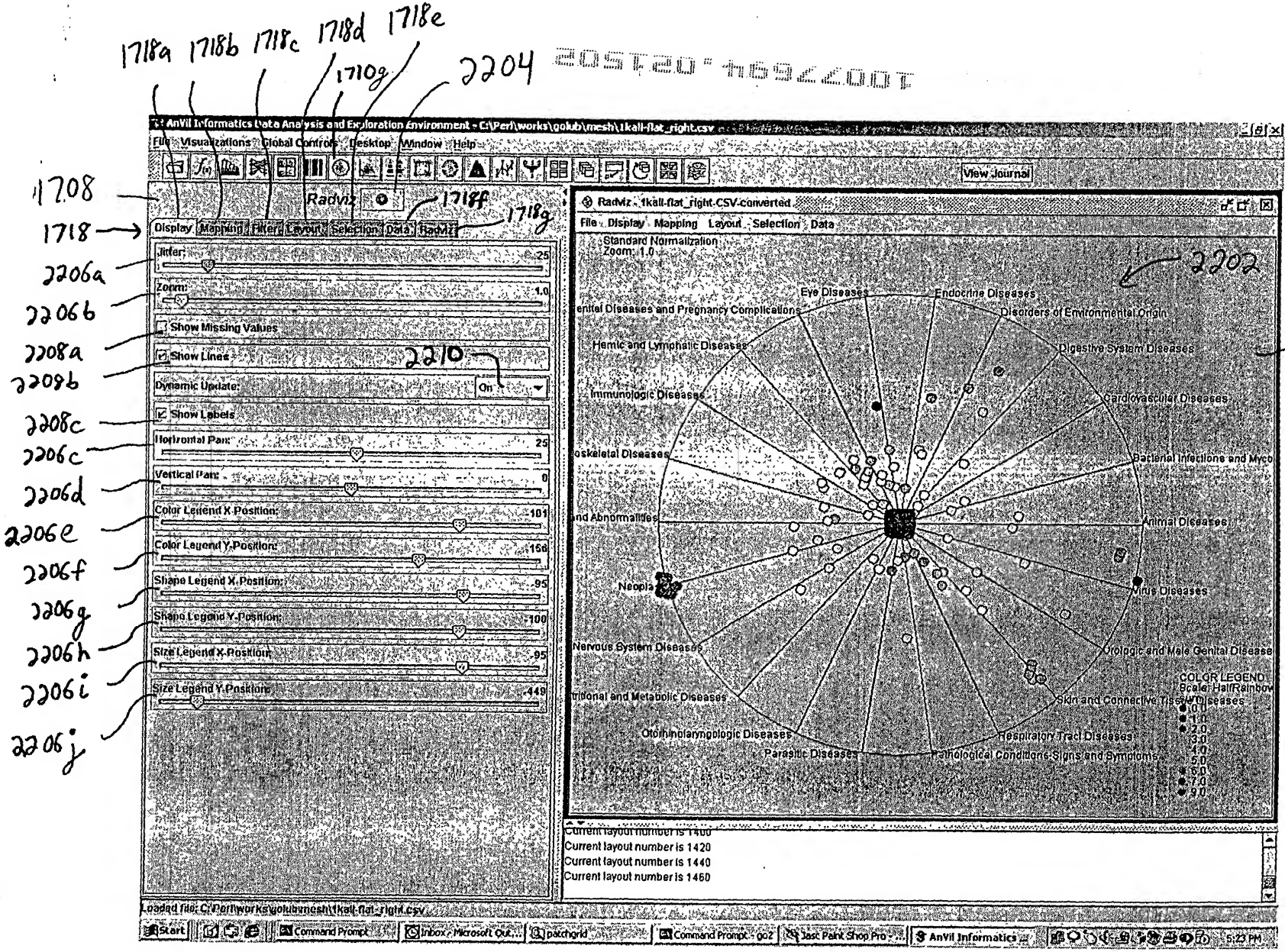


Figure 22

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 52 of 79

22189

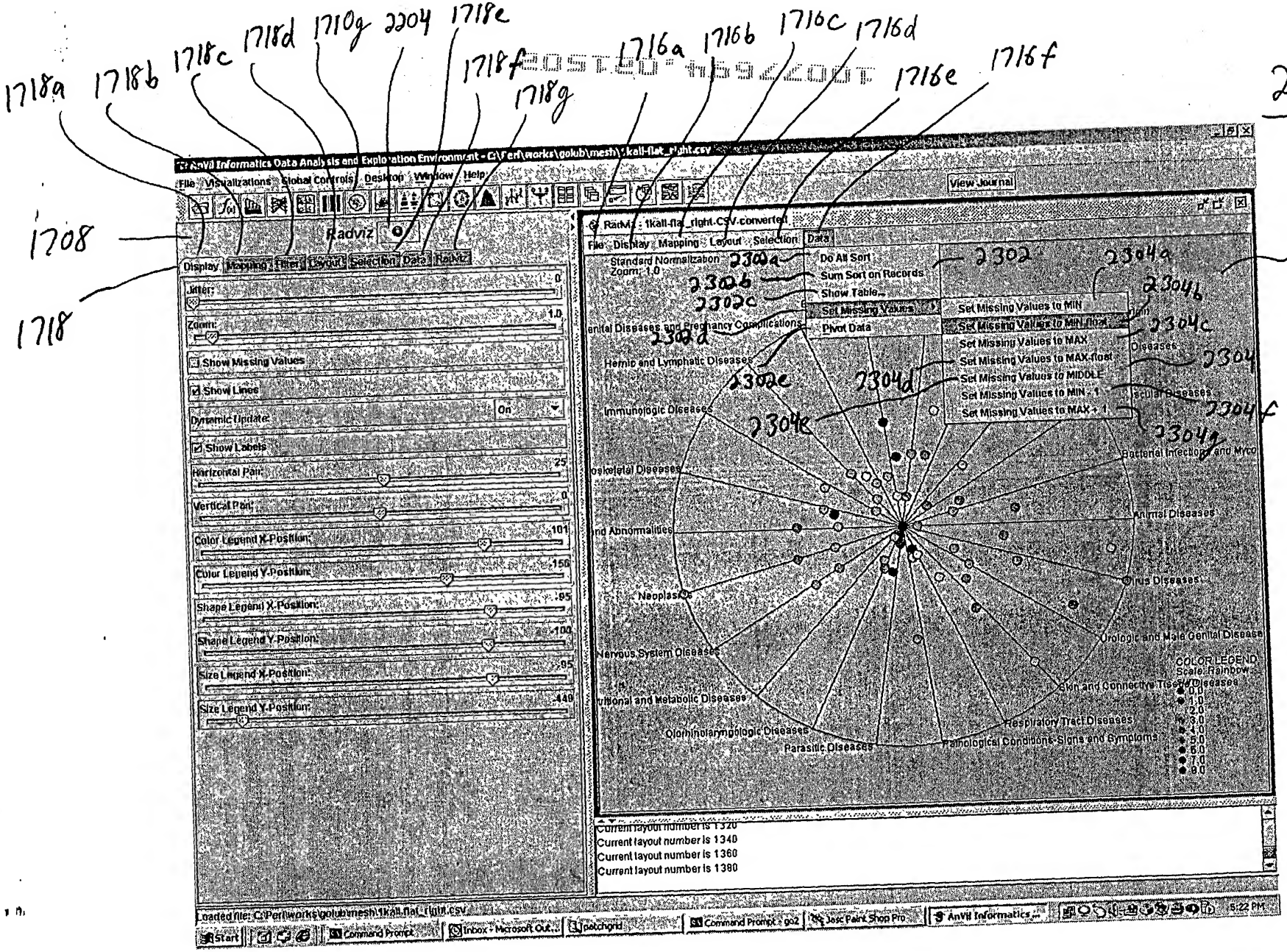


Figure 23

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 53 of 79

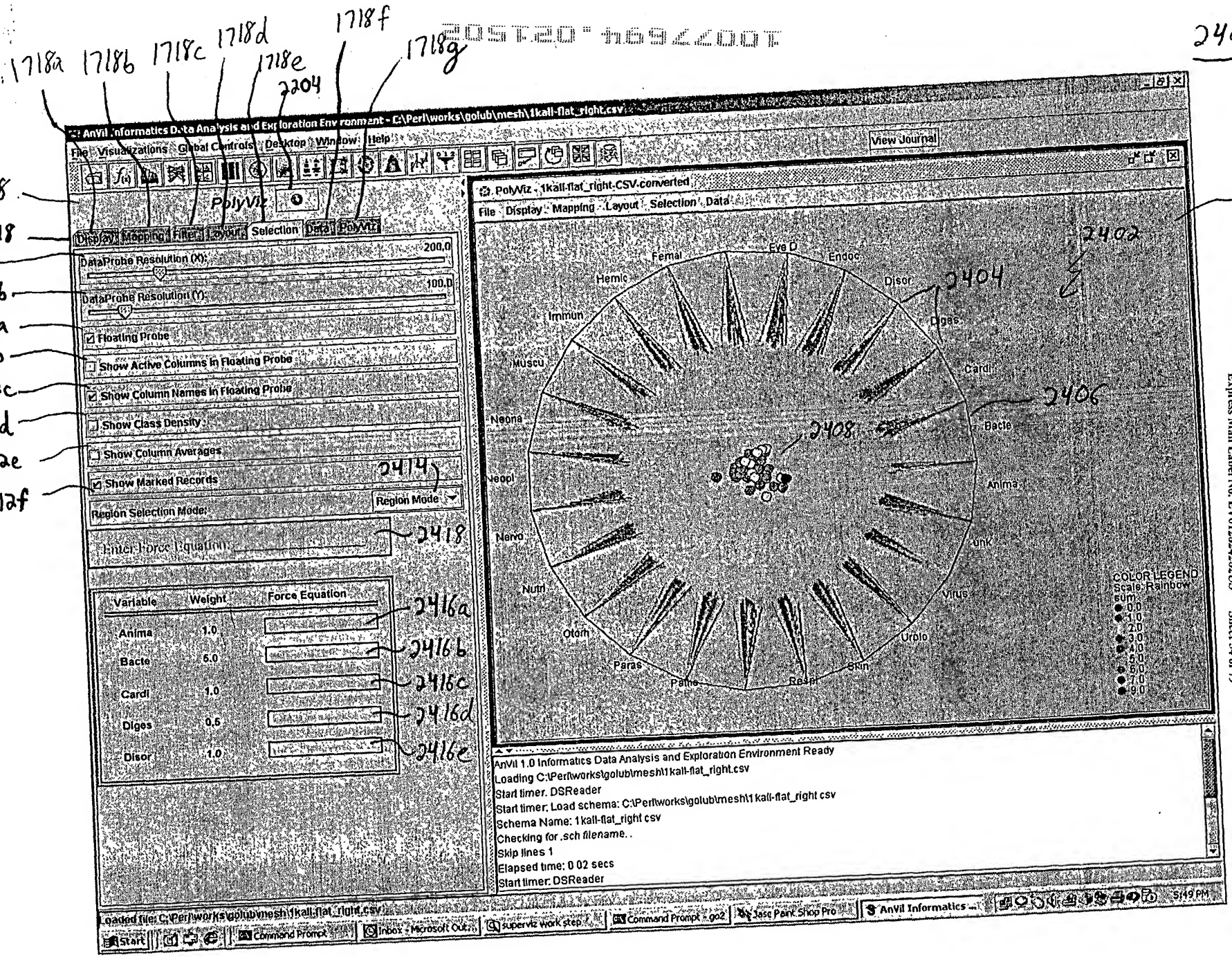


Figure 24

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 54 of 79

22189

2500

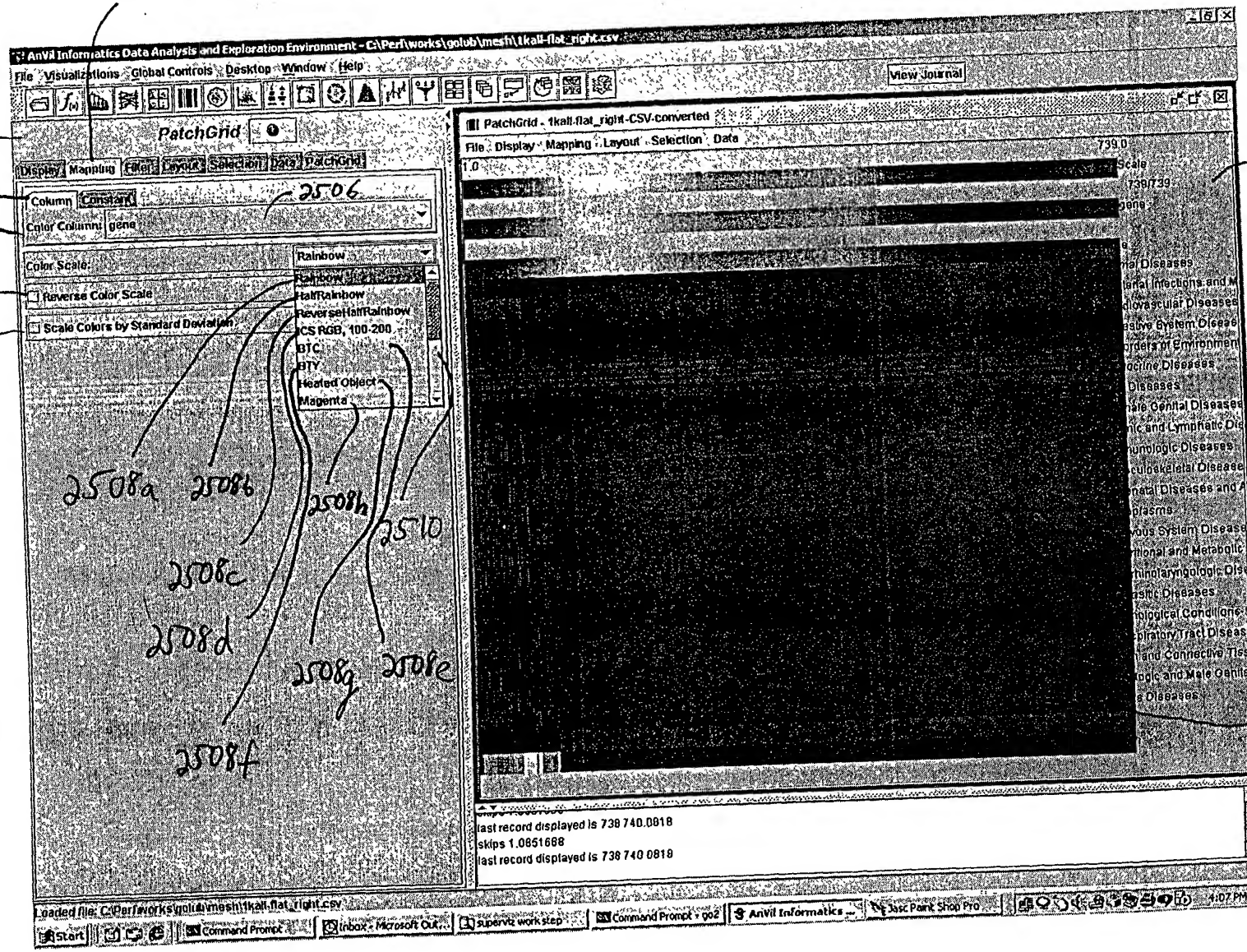


Figure 25

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 55 of 79

22189

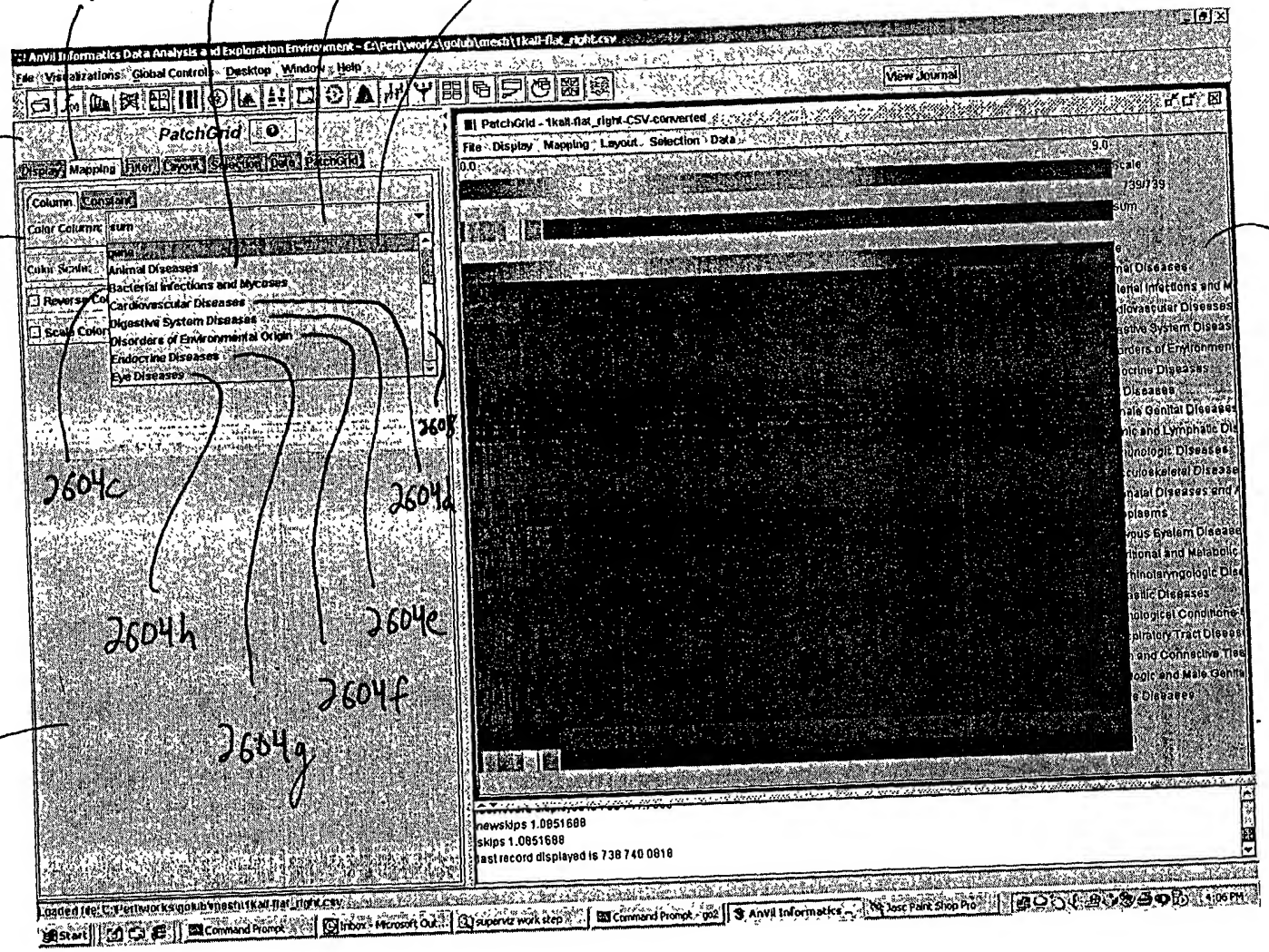


Figure 26

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 56 of 79

221891

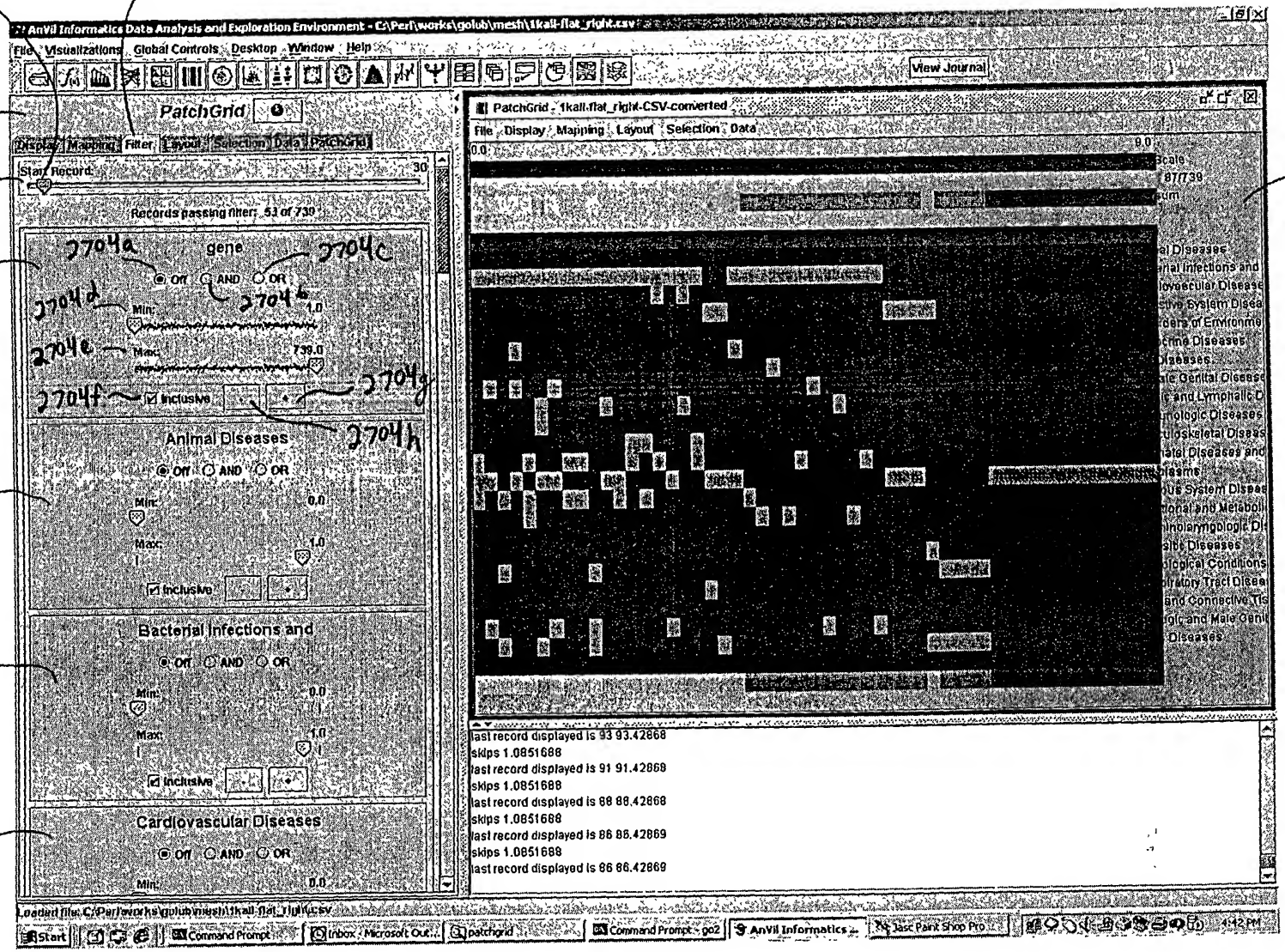

2702a
1718c
2702
1708
2702
2704
2704a
2704c
2704b
2704d
2704e
2704f
2704g
2704h
1706
2706
2708
2710
AnVil Informatics Data Analysis and Exploration Environment - C:\Perl\works\golub\mesh\1kall-flat_right.csv
File Visualizations Global Controls Desktop Window Help
View Journal
PatchGrid
Display Mapping Filter Layout Selection Data PatchGrid
Start Record: 30
Records passing filter: 53 of 730
gene
Off AND OR
Min: 1.0
Max: 739.0
Inclusive
Animal Diseases
Off AND OR
Min: 0.0
Max: 1.0
Inclusive
Bacterial Infections and
Off AND OR
Min: 0.0
Max: 1.0
Inclusive
Cardiovascular Diseases
Off AND OR
Min: 0.0
PatchGrid: 1kall-flat_right-CSV-converted
File Display Mapping Layout Selection Data
last record displayed is 93 93.42868
skips 1.0851688
last record displayed is 91 91.42868
skips 1.0851688
last record displayed is 88 88.42868
skips 1.0851688
last record displayed is 86 86.42869
skips 1.0851688
last record displayed is 86 86.42869
Loaded file C:/Perl/works/golub/mesh/1kall-flat_right.csv
Start Command Prompt Inbox - Microsoft Out... patchgrid Command Prompt - go2 AnVil Informatics ... Jasc Paint Shop Pro ... 4:42 PM

Figure 27

2800

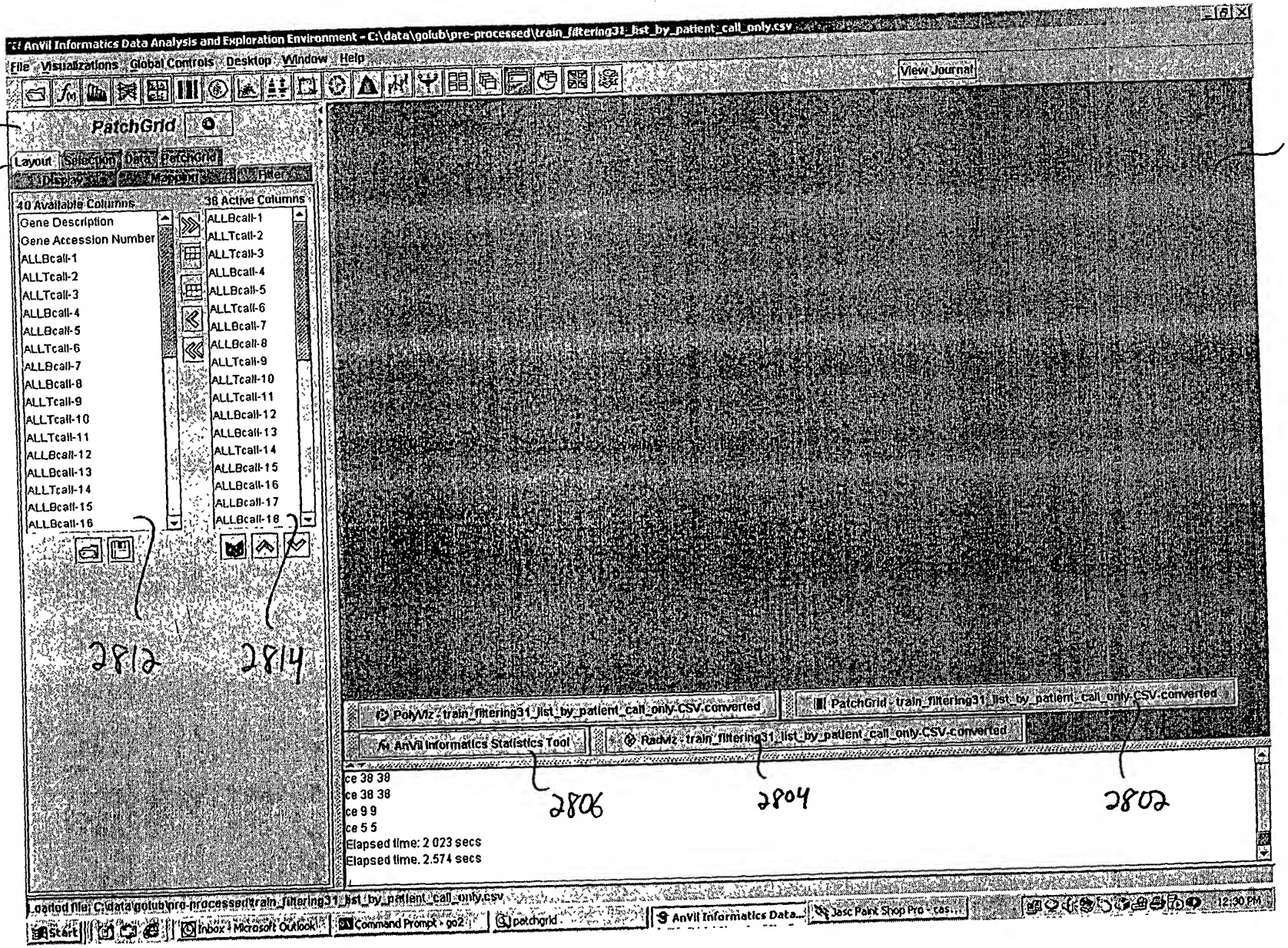


Figure 28

Title: Method And System For Data Analysis
Inventor(s). Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 58 of 79

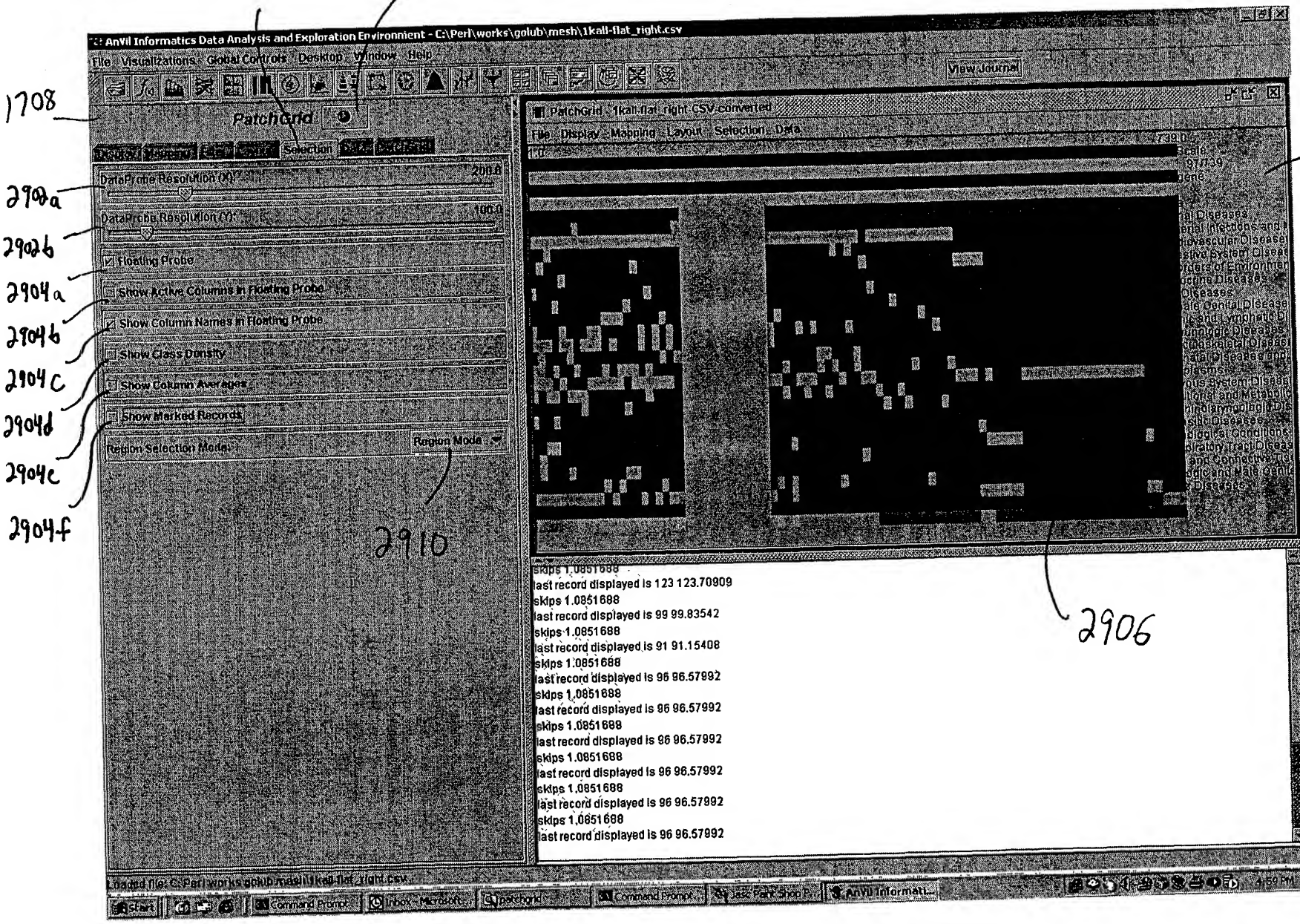


Figure 29

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 59 of 79

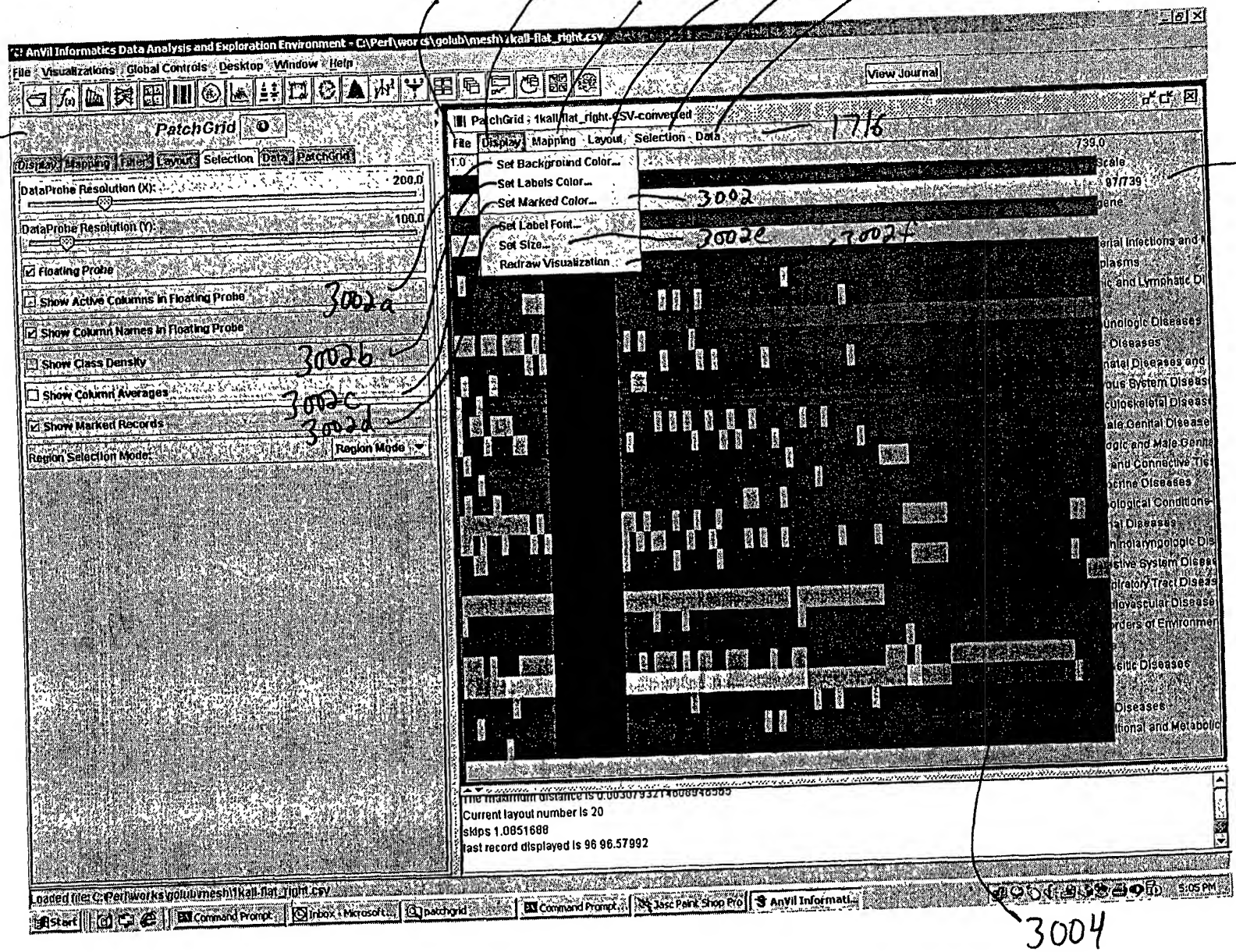


Figure 30

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824382US Sheet 60 of 79

3100

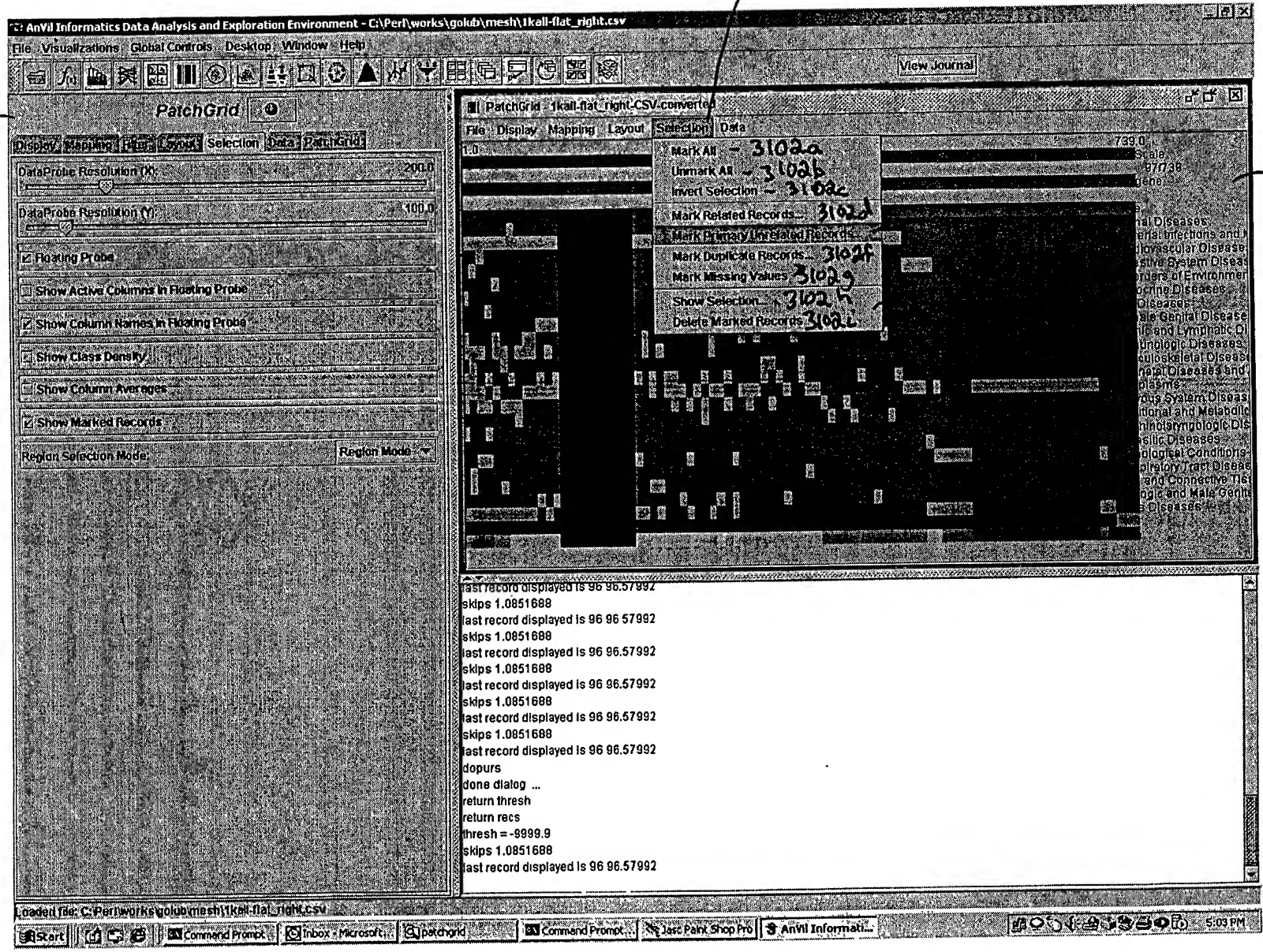


Figure 31

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 61 of 79

3200

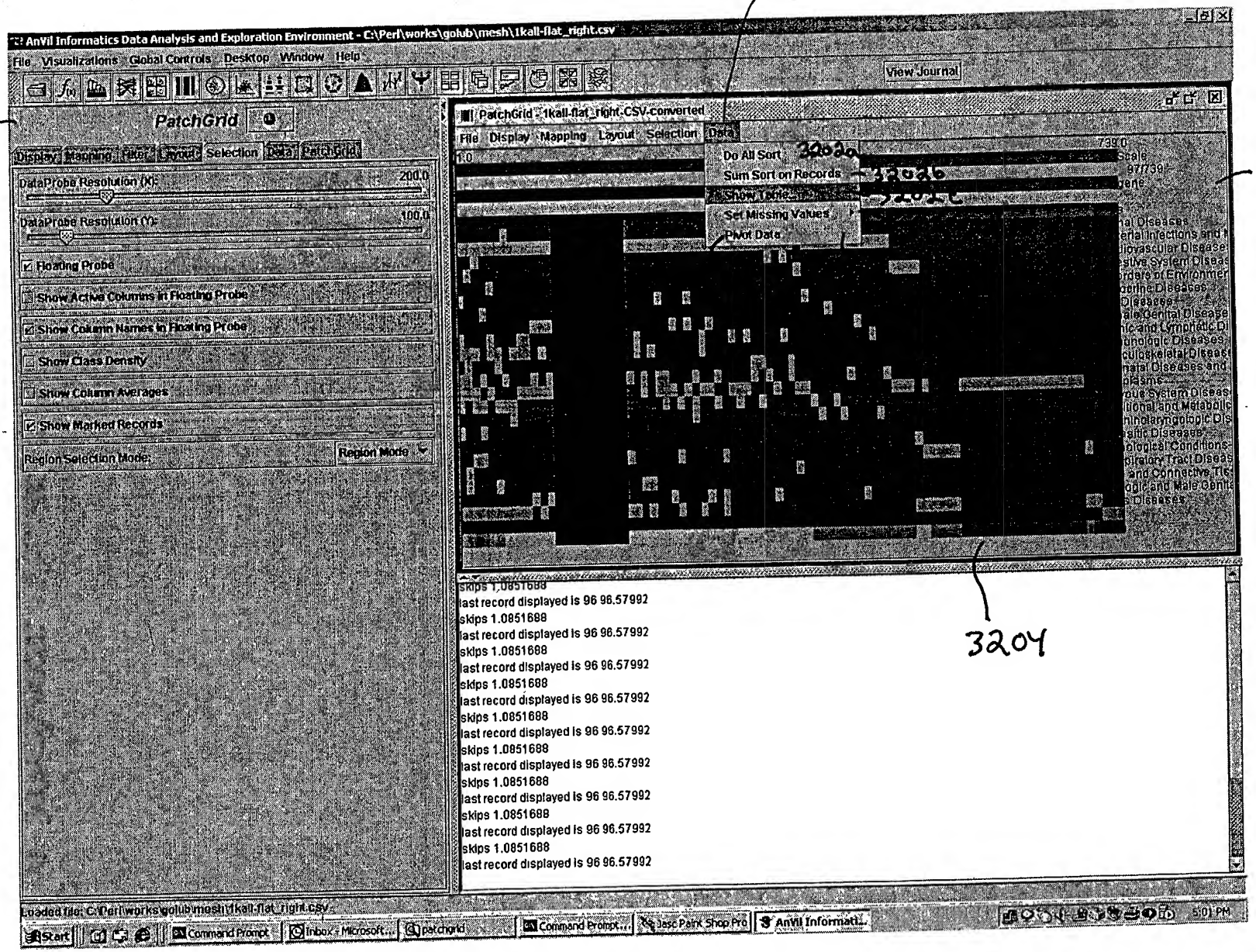
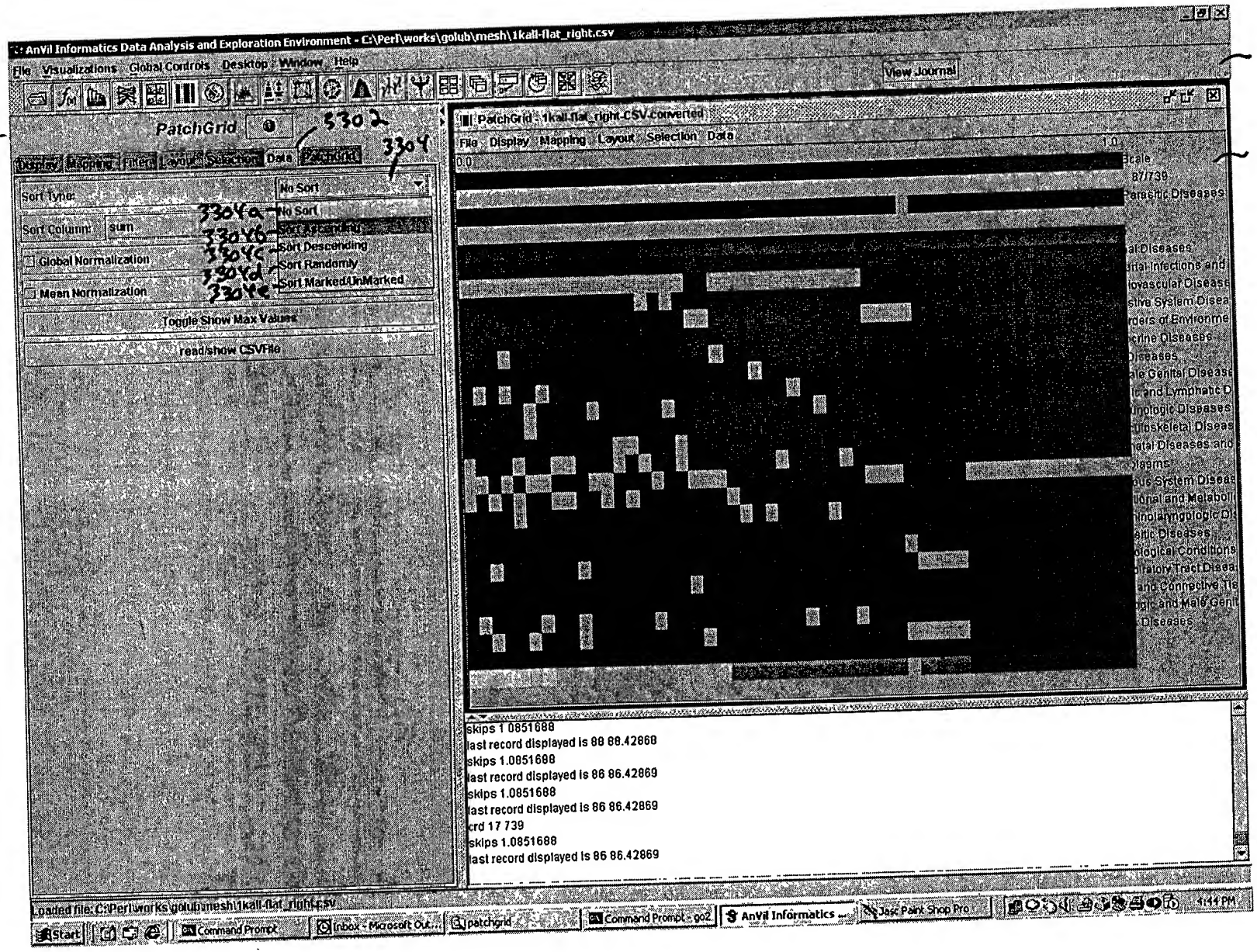


Figure 32

Title: Method And System For Data Analysis
Inventor(s). Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 62 of 79

20ST20 th9c2007



1708

5302

3304

3304a

3304b

3304c

3304d

3304e

1706

Figure 33

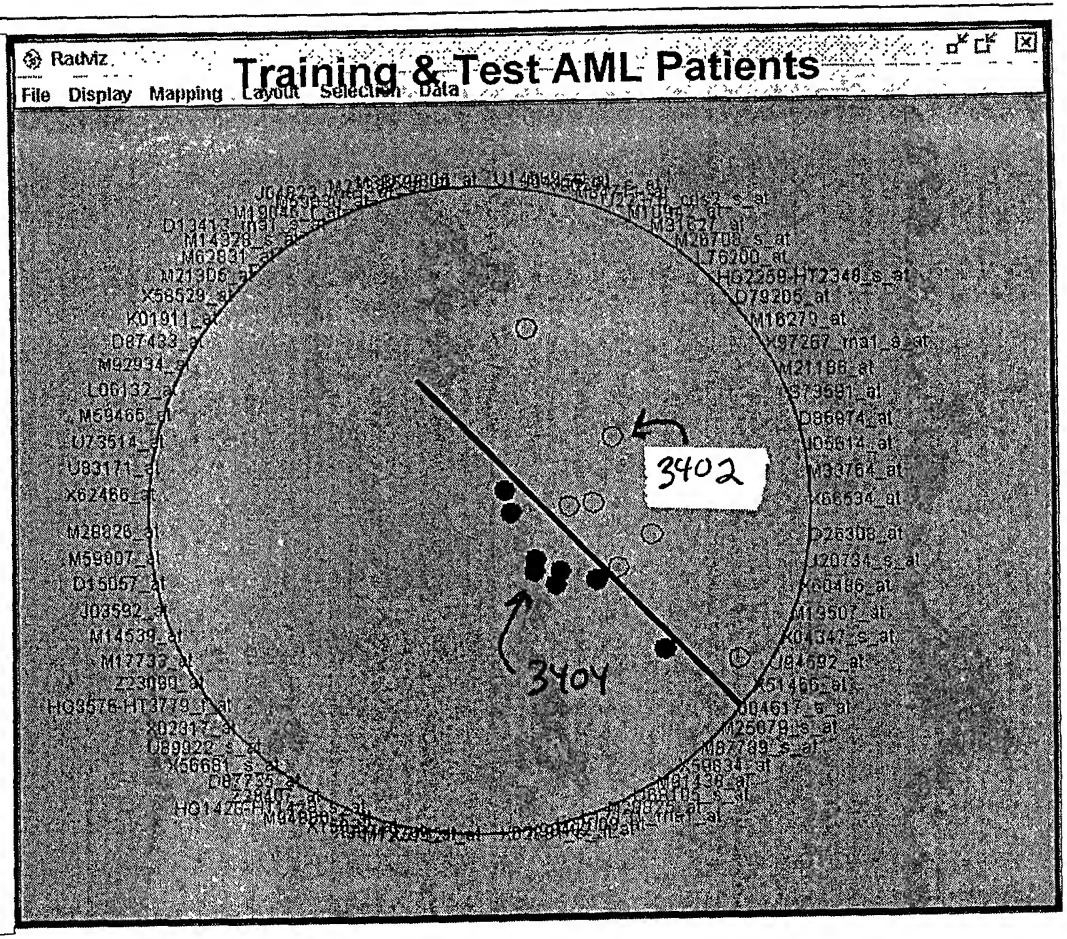
Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 63 of 79

20ST20 "H6927007

Treatment Outcome Prediction

Chemotherapy

- 76 Gene Predictor Set
  - **Success**
  - **Failure**
- Boundary



Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 64 of 79

Figure 34

20ST20 "16SC00F

Reduced Genes

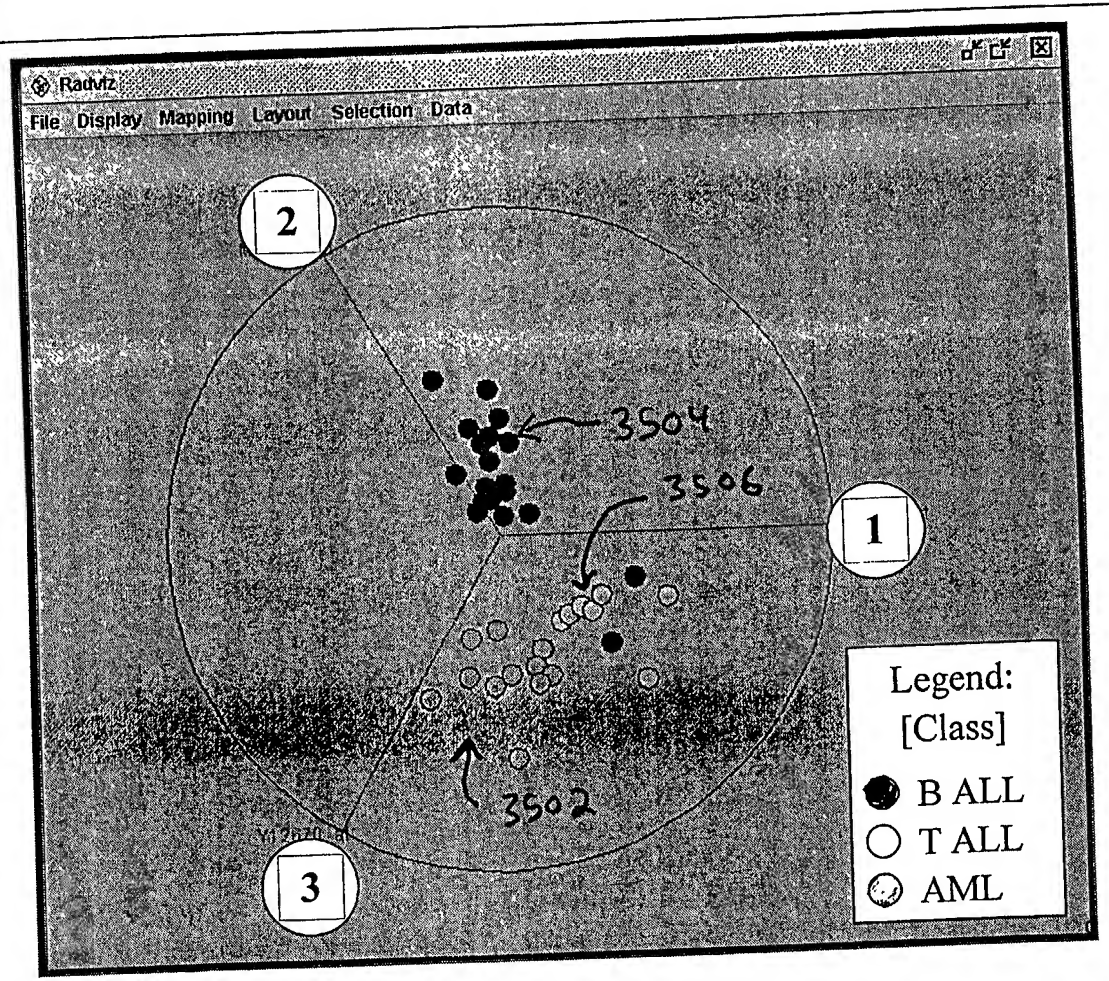


Figure 35

Title: Method And System For Data Analysis
Inventor(s). Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 65 of 79

2259884

20ST20 "h69c2001"

Chemical Structure

Benzodiazepines

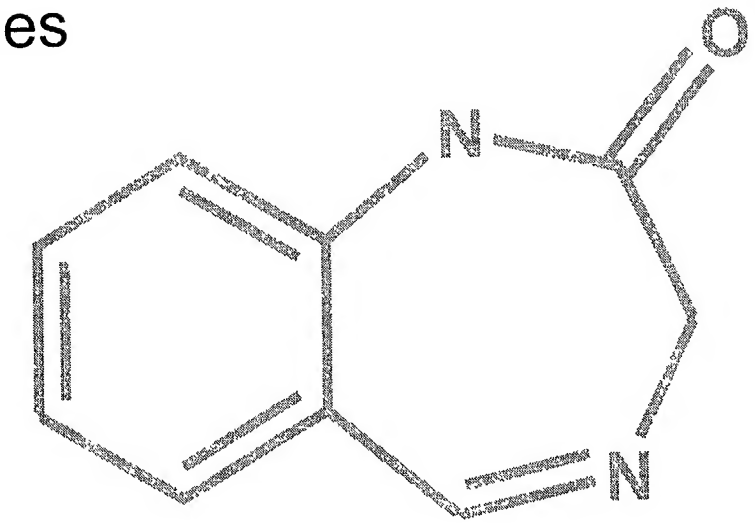


Figure 36

20ST20 "h69z200T

Radial Visualization Showing R_3 Attributes

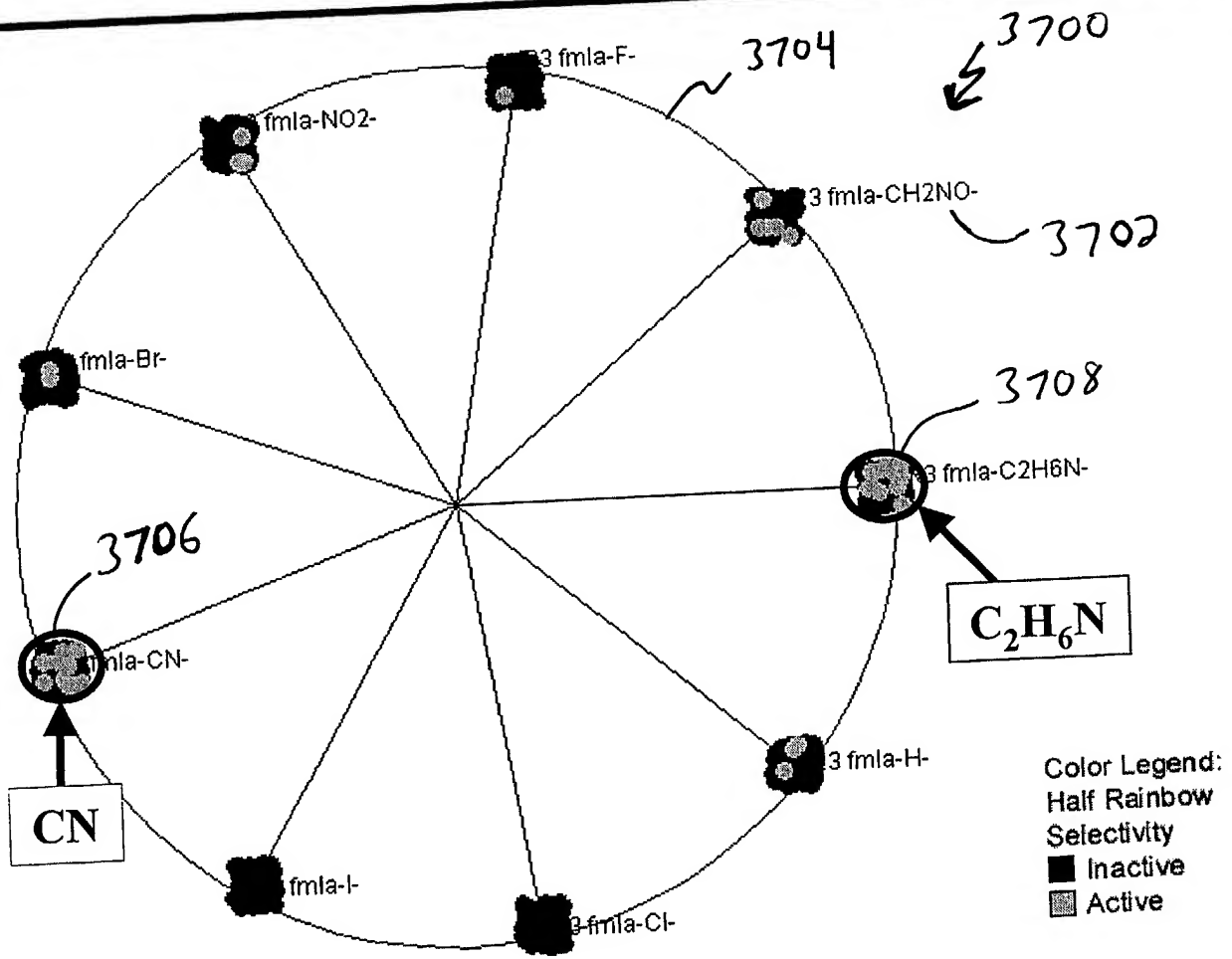
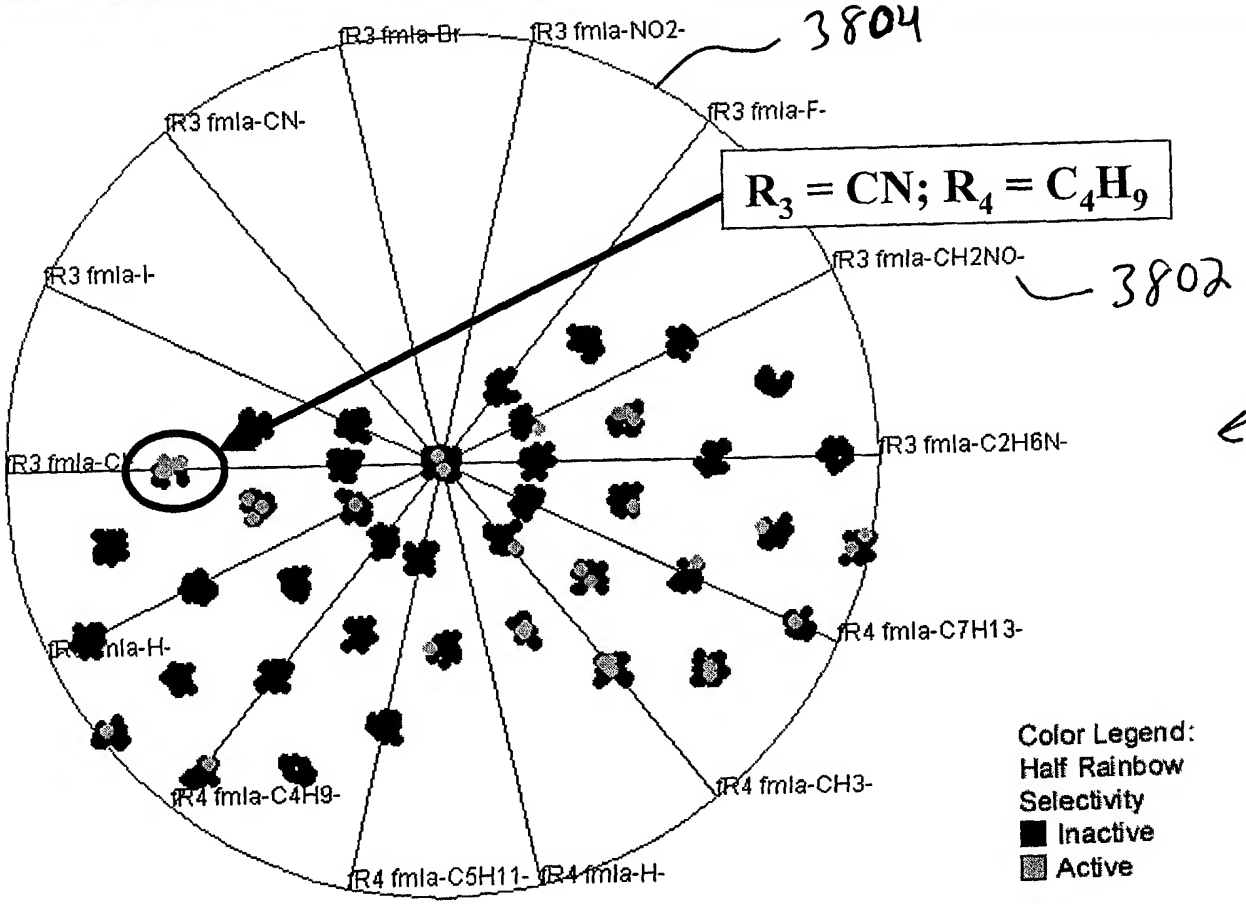


Figure 37

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 67 of 79

20520 "h69c00t

Radial Visualization Showing R_3 & R_4 Attributes



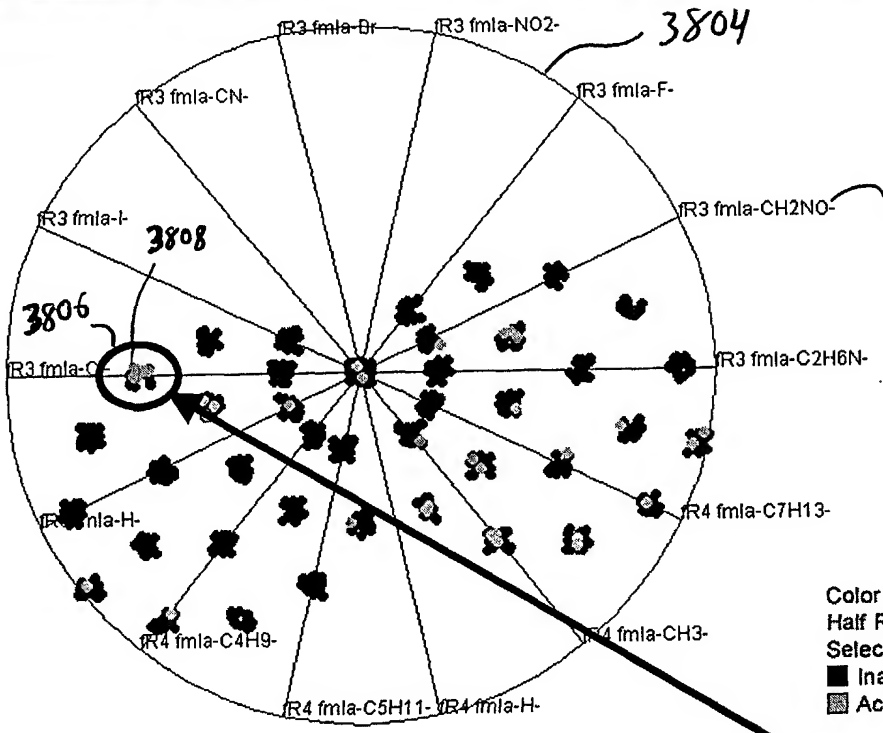
Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 68 of 79

Figure 38

2273661

20020 h69720Ot

Radial Visualization Showing R_3 & R_4 Attributes



Record	LC/MS	Assay 1	Assay 2	Kd 1	Kd 2	R1	R2	R3	R4	S5	Select
937	pass	0.087	355.5	591.7	9.51	C5H5-	C6H11-	CN-	C4H9-	3.00	good
937	pass	0.01	456.4	0.87	0.11	CH3-	C6H11-	CN-	C4H9-	4.43	good
937	pass	0.007	358.8	44.7	3.89	H-	C6H5-	CN-	C4H9-	5.58	good
937	pass	0.002	495.8	78.35	57.52	C9H8N-	C6H11-	CN-	C4H9-	6.44	good

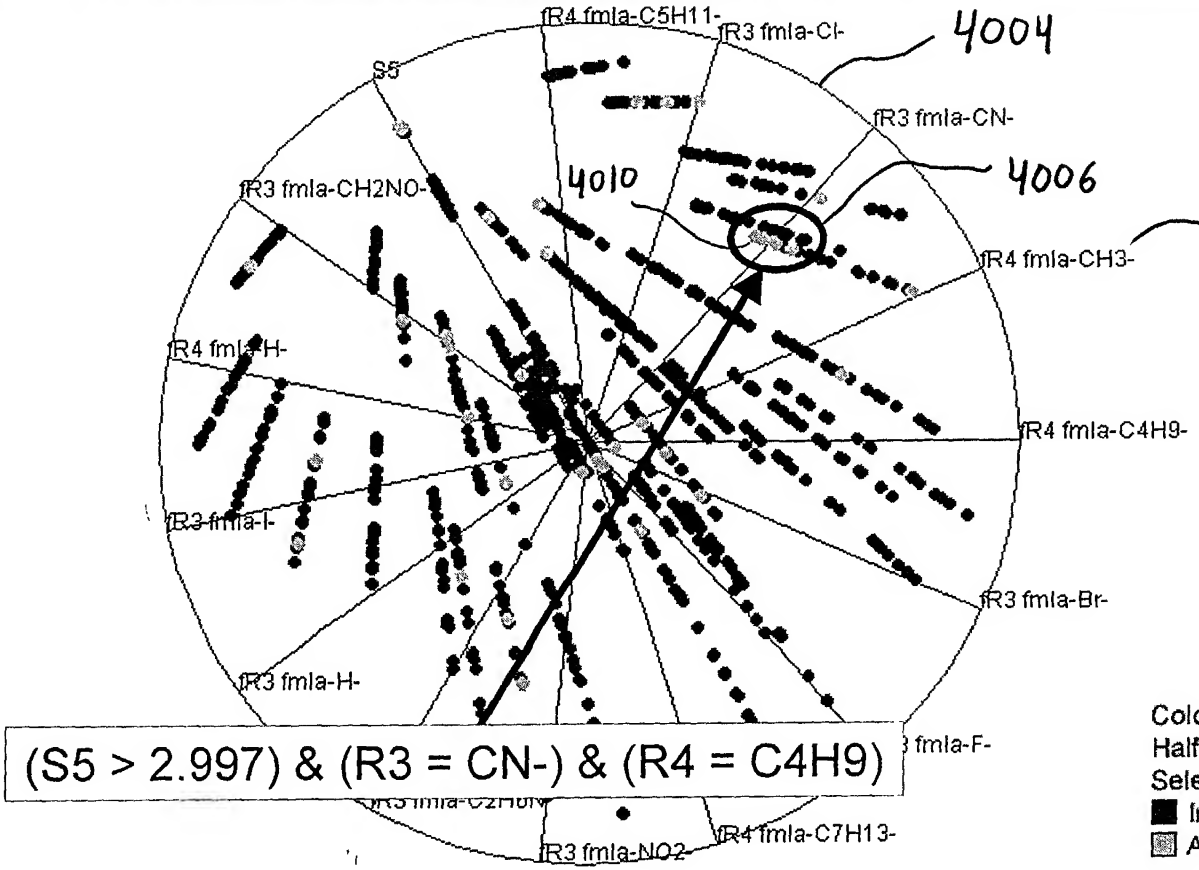
3810

Figure 39

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 69 of 79

20STQP * 16927001

Radial Visualization Showing R_3, R_4 & S_5 Attributes



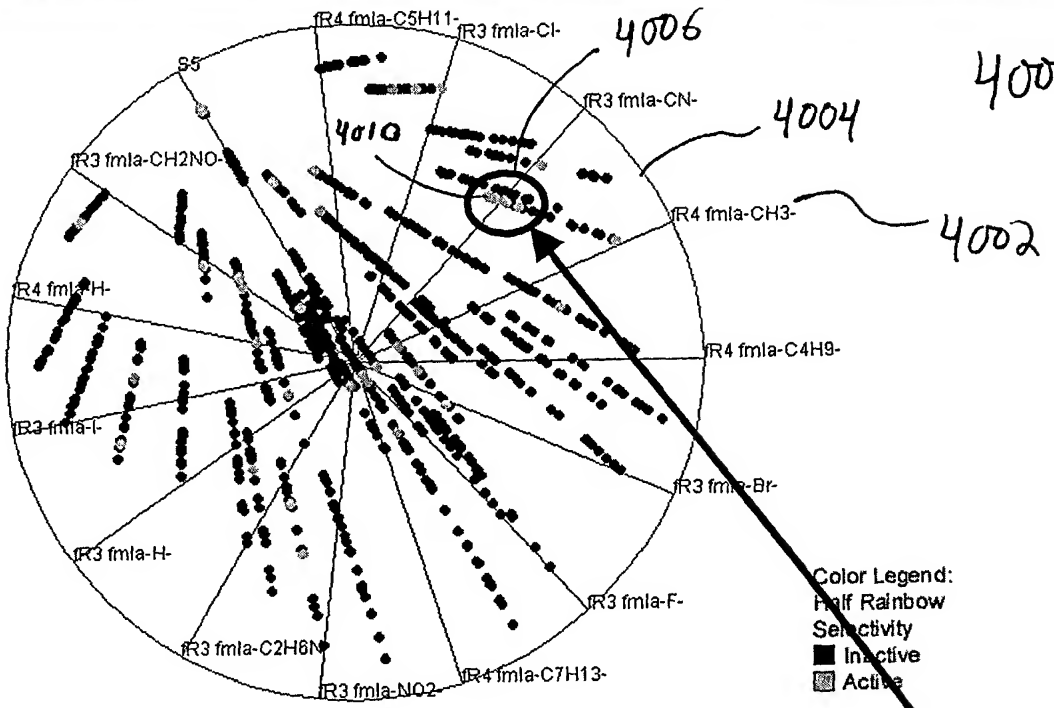
Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 70 of 79

Figure 40

20020-1692600

Radial Visualization

Showing R_3, R_4 & S_5 Attributes



Record	LC/MS	Assay 1	Assay 2	Kd 1	Kd 2	R1	R2	R3	R4	S5	Select.
937	pass	0.087	355.5	591.7	9.51	C5H5-	C6H11-	CN-	C4H9-	3.00	good
951	pass	0.01	456.4	0.87	0.11	CH3-	C6H11-	CN-	C4H9-	4.43	good
953	pass	0.007	358.8	44.7	3.89	H-	C6H5-	CN-	C4H9-	5.58	good
956	pass	0.002	495.8	78.35	57.52	C9H8N-	C6H11-	CN-	C4H9-	6.44	good
958	pass	0.001	305.8	0.14	7.6	H-	C6H11-	CN-	C4H9-	7.08	good

4008

Figure 41

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 71 of 79

20STEP-H692.DOC

Pearson Cross Correlations

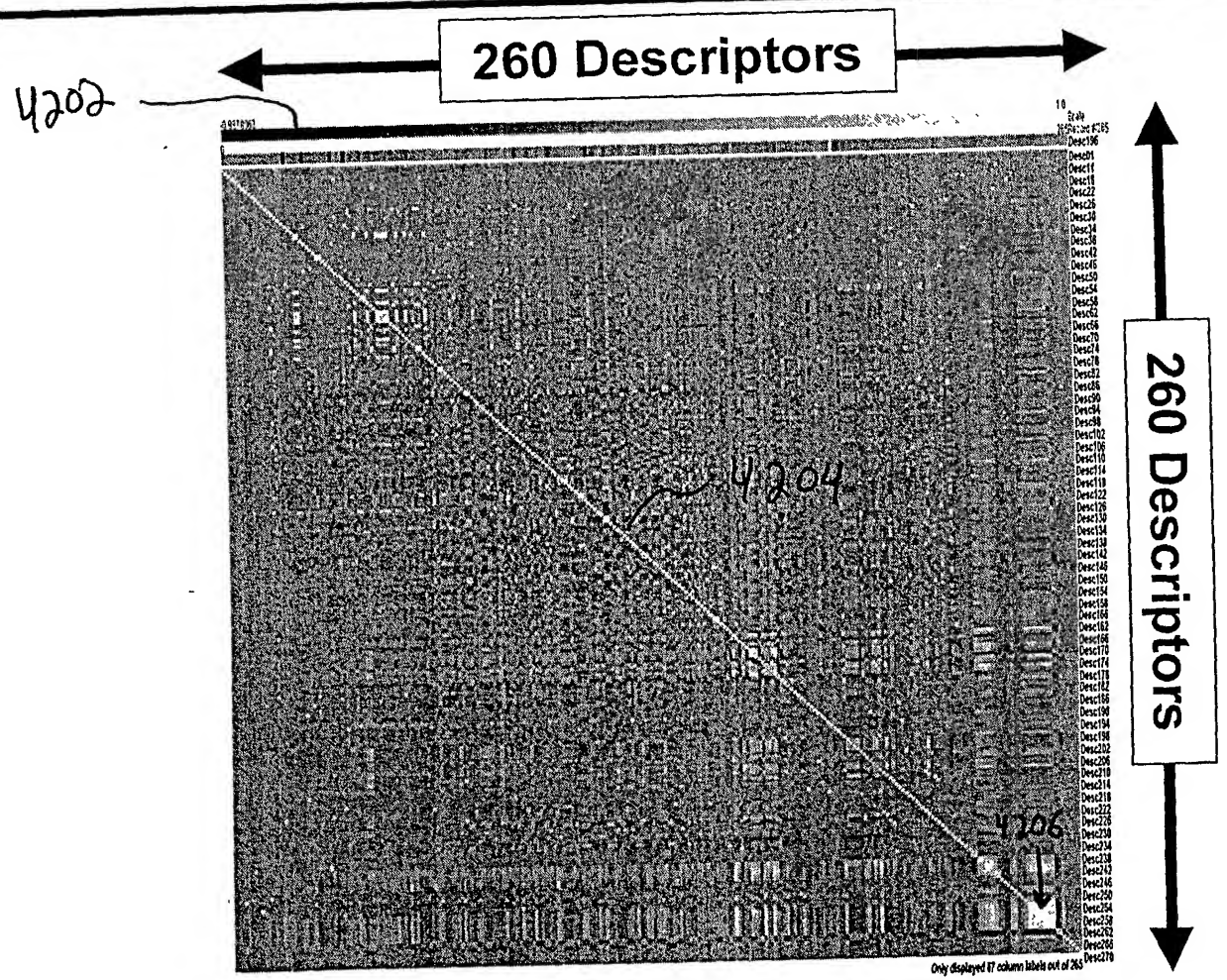


Figure 42

20STP20 h692200T

ISIS Keys

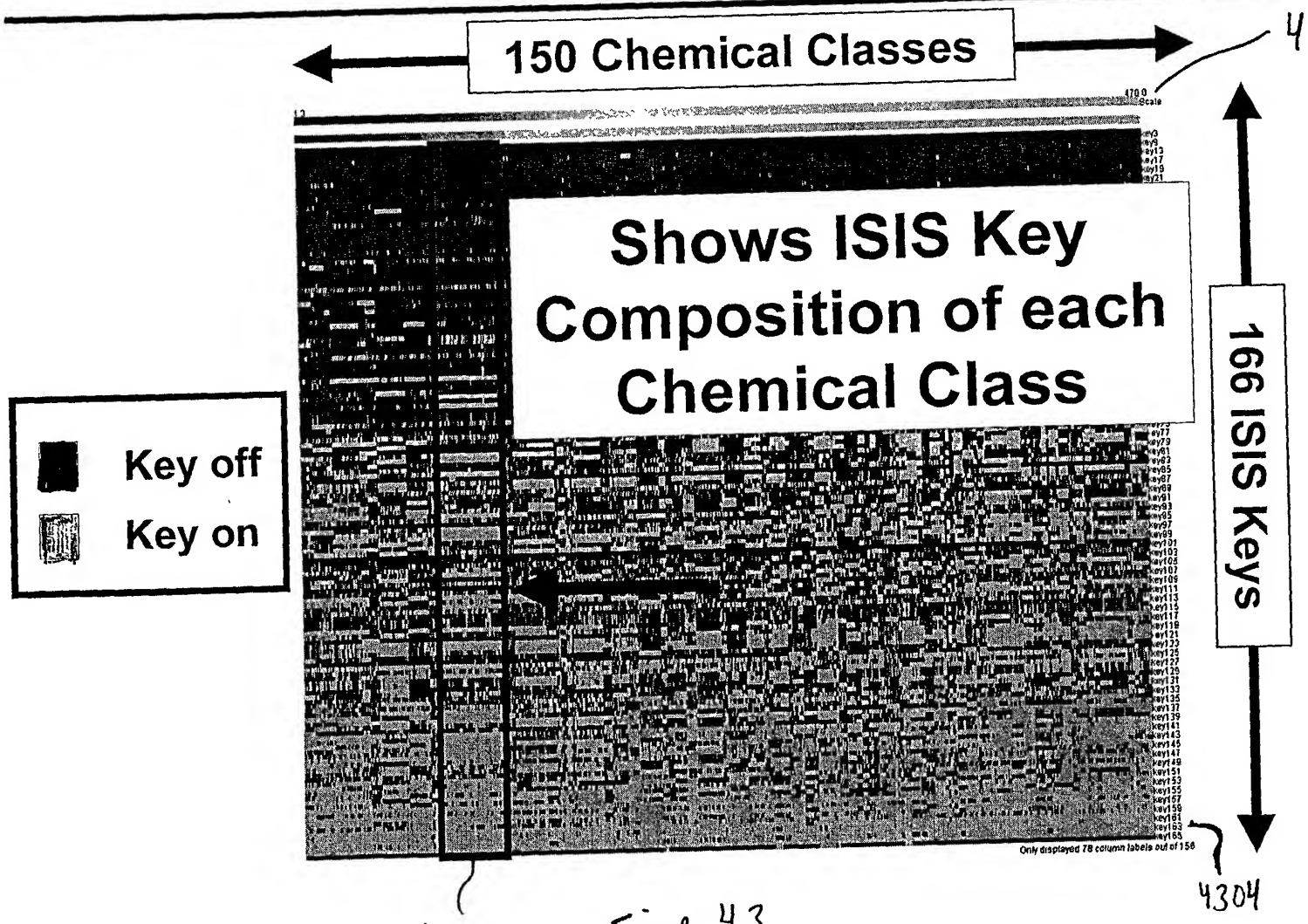


Figure 43

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 73 of 79

2051,20 "4692007"

ISIS Keys

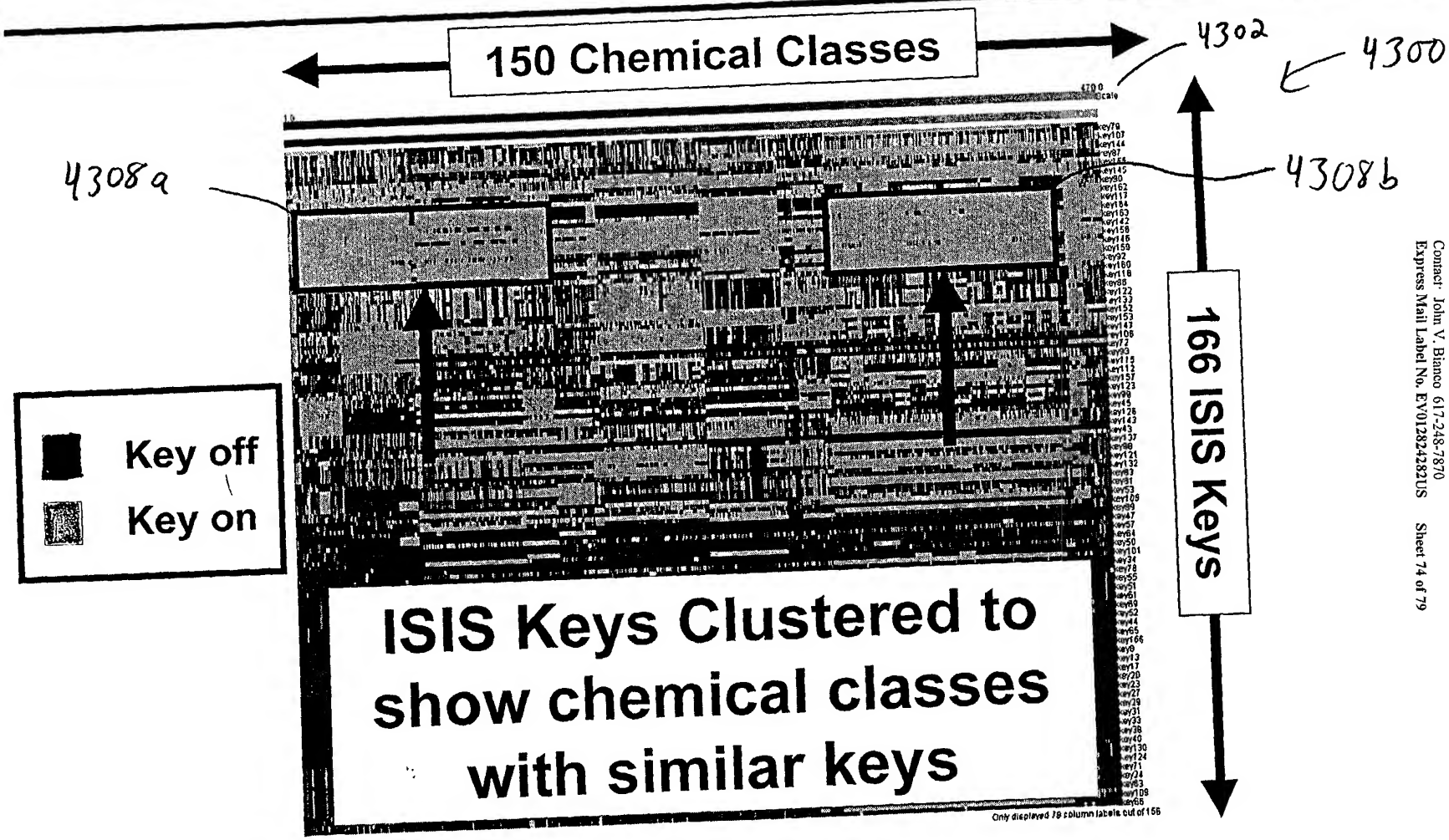
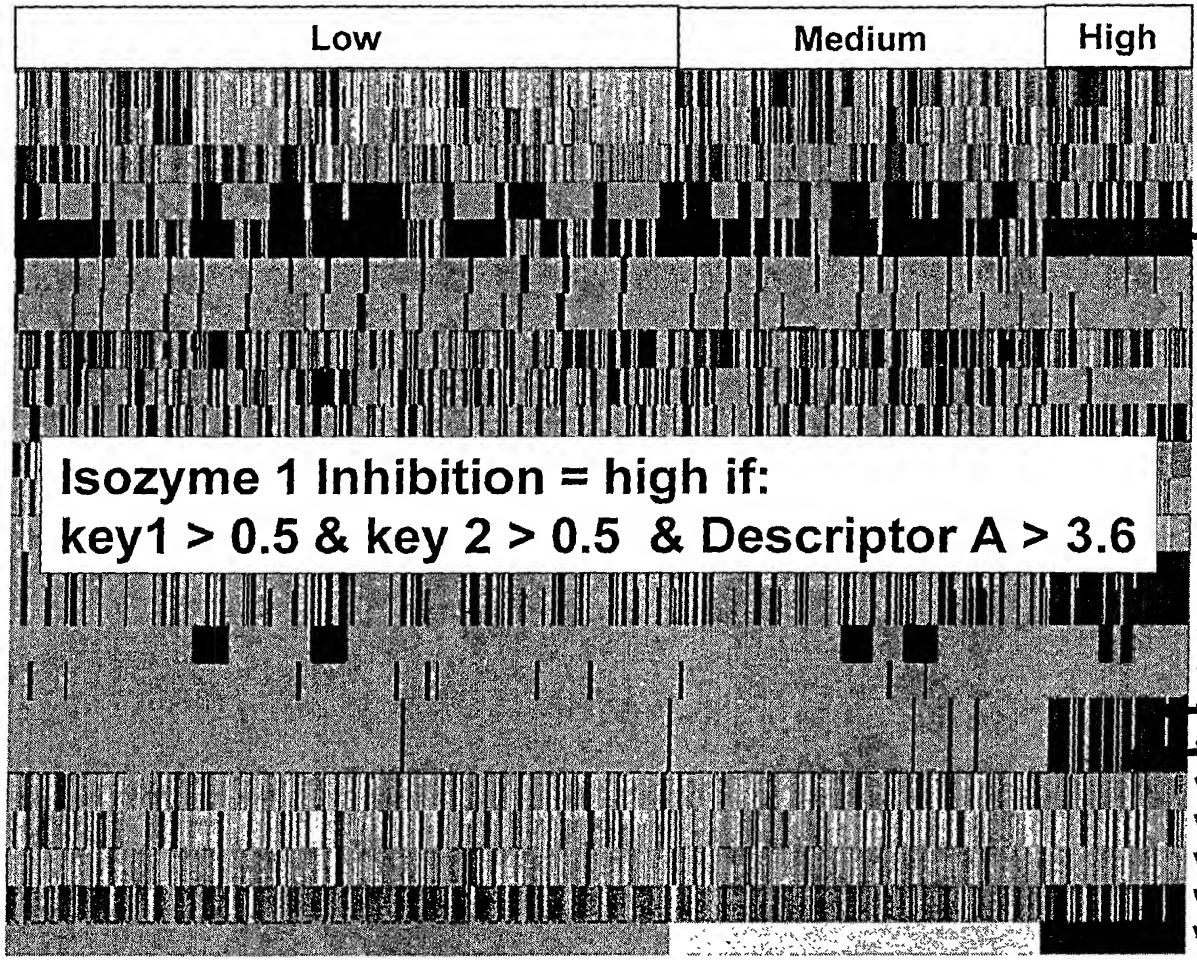


Figure 44

20 5 F 20 - H 5 9 7 2 0 0 T

Association Rules - Visually



45-00

Figure 45

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 75 of 79

2273661

20st20 4692004

Exploratory Overview

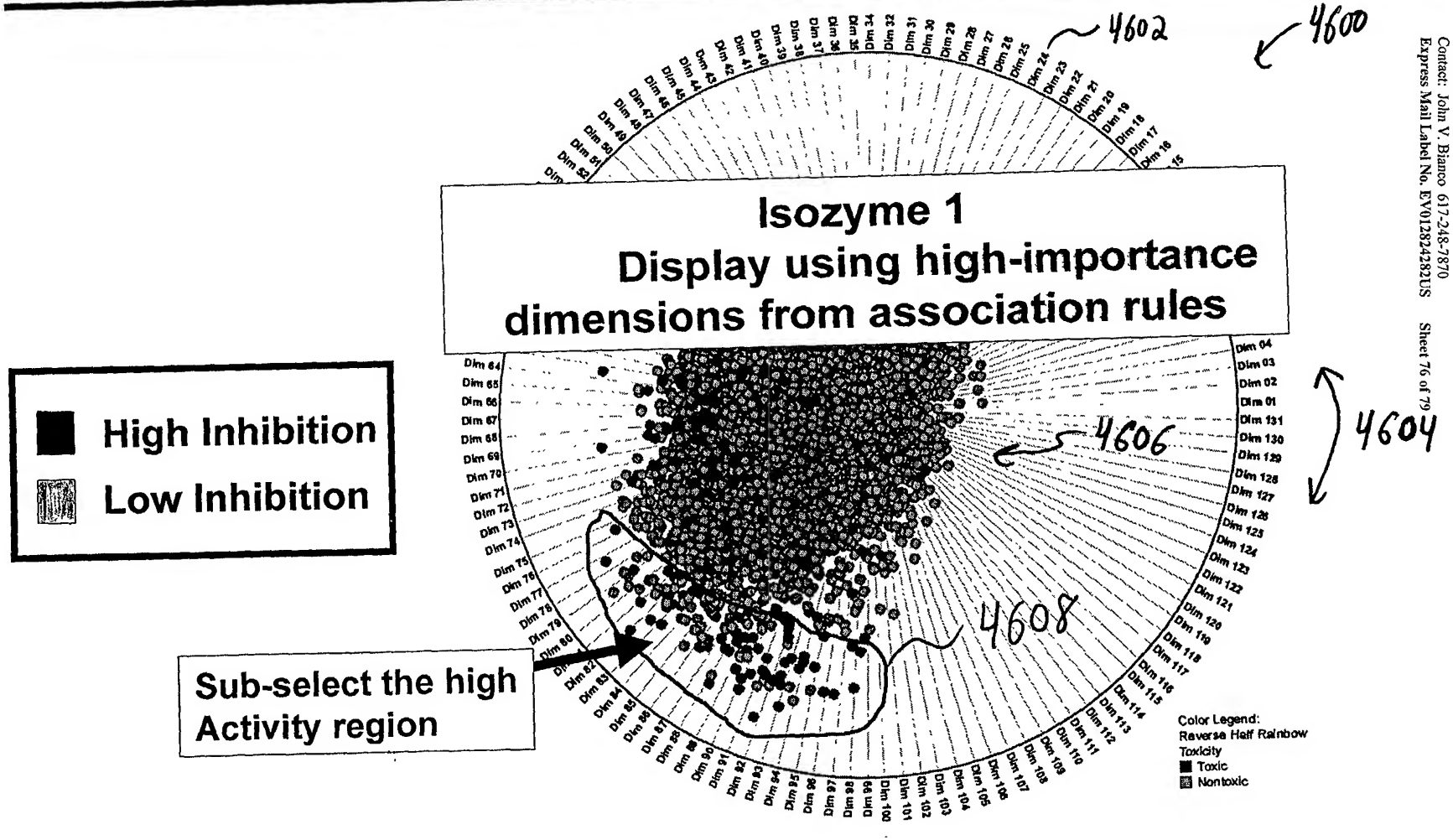


Figure 46

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 76 of 79

20ST20 th692c00T

Toxic Distinction

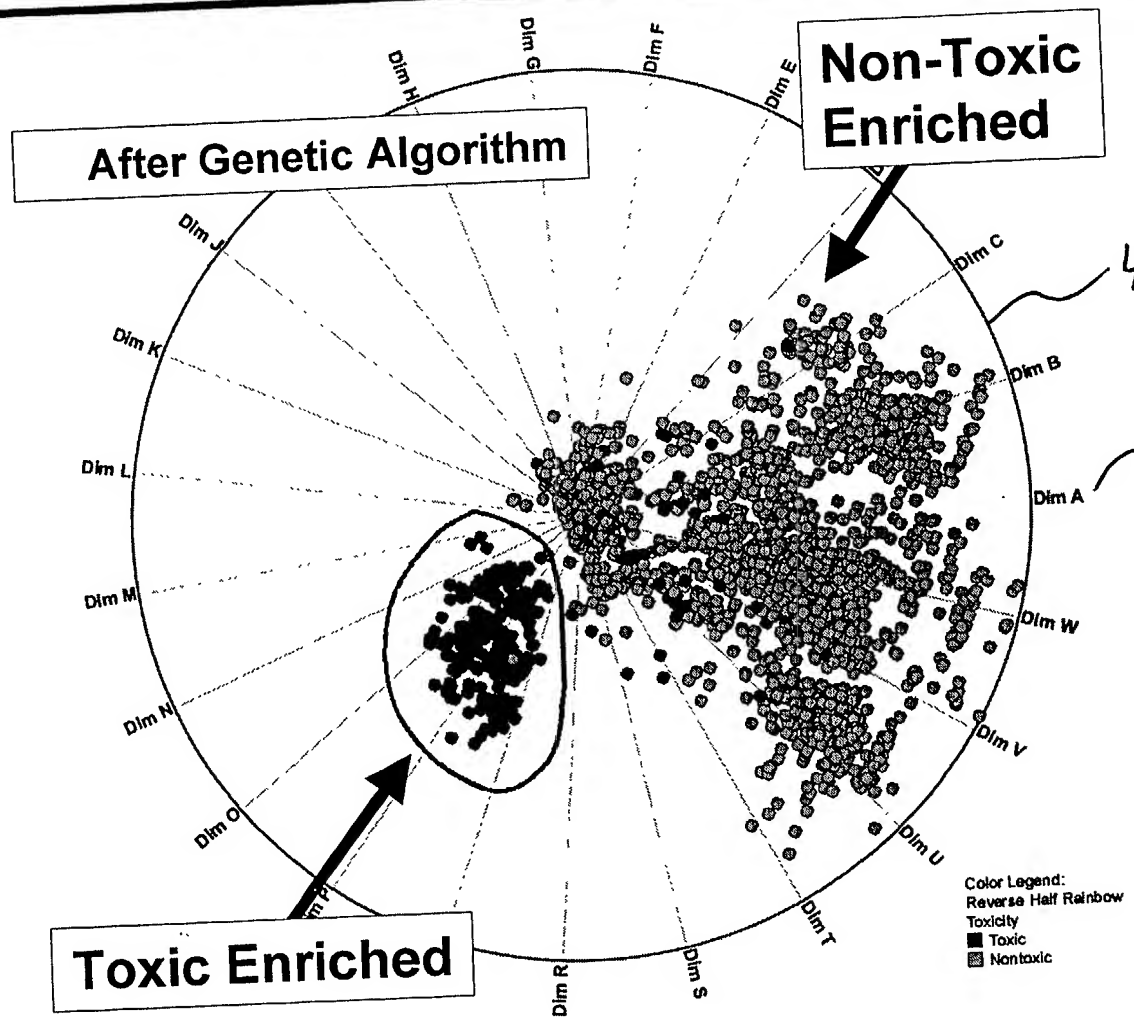


Figure 47

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V Bianco 617-248-7870
Express Mail Label No. EV012824282US Sheet 77 of 79

20020*h892d00f

Toxic Distinction

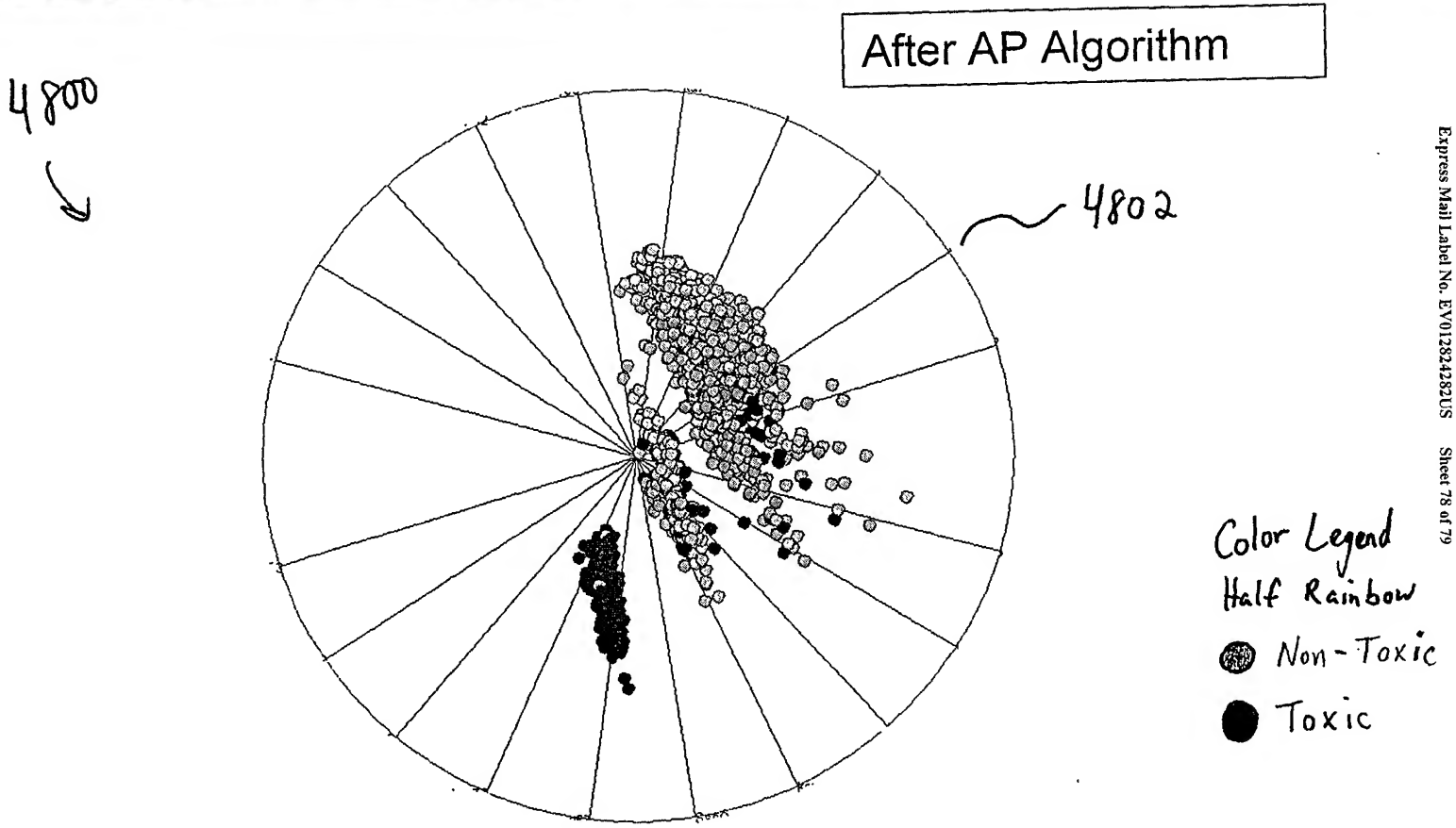


Figure 48

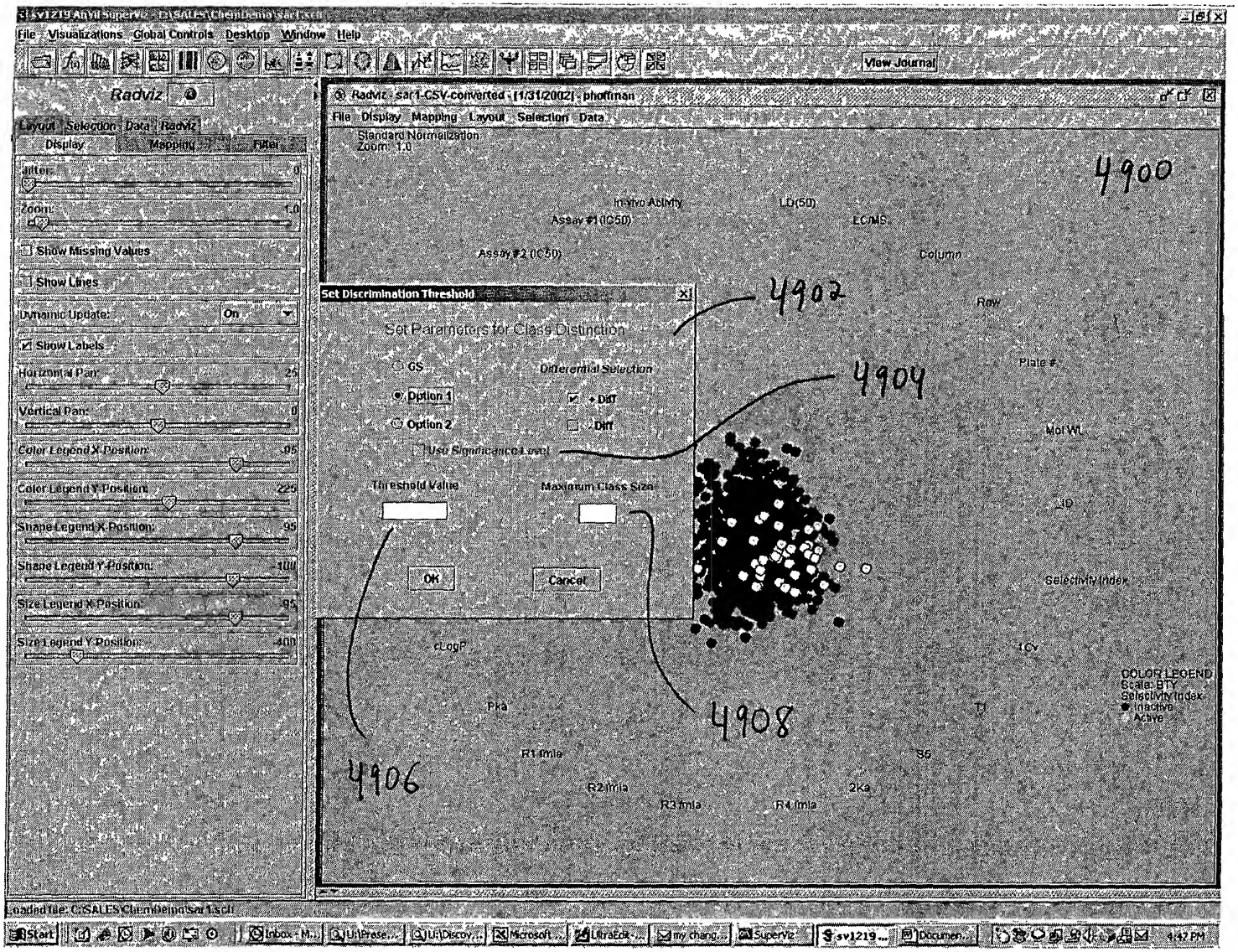


Figure 49

Title: Method And System For Data Analysis
Inventor(s): Gee et al.
Serial No. Not yet assigned
Docket No. ANV-001
Contact: John V. Bianco 617-248-7870
Express Mail Label No. EV012824282US